

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 07:50:49 ; Search time 21.93 Seconds
(without alignments)
31.186 Million cell updates/sec

Title: US-09-856-199-2

Perfect score: 136

Sequence: 1 TGSATXQCKENLPSLCSXQGLRAEENI 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCPTUS_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	34.6	1384	4	US-08-976-255-11
2	45	33.1	659	4	US-09-228-986-75
3	43	31.6	326	2	US-08-997-080-43
4	43	31.6	326	2	US-08-997-362-43
5	43	31.6	326	3	US-08-873-970-43
6	43	31.6	326	4	US-09-095-855-43
7	43	31.6	326	4	US-08-705-347A-43
8	43	31.6	326	4	US-09-324-542-43
9	42	30.9	251	2	US-08-460-529B-2
10	42	30.9	302	3	US-08-831-132-2
11	42	30.9	302	4	US-09-416-150-2
12	42	30.9	1479	3	US-08-840-062-2
13	42	30.9	1479	3	US-08-840-062-4
14	41	30.1	73	2	US-08-716-188-1
15	41	30.1	96	1	US-08-480-449-24
16	41	30.1	96	2	US-08-660-542-24
17	41	30.1	96	4	US-08-679-493A-162
18	41	30.1	96	4	US-08-479-603-24
19	41	30.1	900	2	US-08-813-940-4
20	40.5	29.8	1156	3	US-08-996-083-1
21	40.5	29.8	1156	4	US-09-429-516-1
22	40.5	29.8	1156	4	US-09-429-516-3
23	40	29.4	75	4	US-09-227-357-299
24	40	29.4	113	2	US-08-256-568B-92
25	40	29.4	113	4	US-09-038-369B-92
26	40	29.4	548	3	US-08-688-988-41
27	40	29.4	3066	4	US-08-952-127-12

28	39.5	29.0	1449	3	US-08-840-062-6	Sequence 6, Appli
29	39	28.7	184	4	US-09-452-817-1	Sequence 1, Appli
30	39	28.7	207	1	US-08-588-163-5	Sequence 5, Appli
31	39	28.7	207	2	US-09-111-070-5	Sequence 5, Appli
32	39	28.7	207	4	US-08-849-764C-5	Sequence 5, Appli
33	39	28.7	370	4	US-08-857-076-104	Sequence 104, App
34	39	28.7	463	4	US-09-163-444-2	Sequence 2, Appli
35	39	28.7	539	3	US-08-688-988-42	Sequence 42, Appli
36	39	28.7	1382	2	US-08-737-715-2	Sequence 2, Appli
37	38.5	28.3	319	2	US-08-683-908-1	Sequence 1, Appli
38	38	27.9	505	1	US-08-220-603A-10	Sequence 10, Appli
39	38	27.9	1463	1	US-08-220-603A-11	Sequence 11, Appli
40	37.5	27.6	287	1	US-07-877-516-3	Sequence 3, Appli
41	37	27.2	4207	4	US-09-068-960-43	Sequence 43, Appli
42	37	27.2	225	4	US-09-040-483-1	Sequence 1, Appli
43	37	27.2	271	4	US-09-181-183-36	Sequence 36, Appli
44	37	27.2	271	4	US-09-277-700-36	Sequence 36, Appli
45	37	27.2	301	3	US-08-829-525-24	Sequence 24, Appli
46	37	27.2	301	4	US-08-609-583A-24	Sequence 24, Appli
47	37	27.2	301	4	US-08-937-399-24	Sequence 24, Appli
48	37	27.2	321	4	US-08-915-795-9	Sequence 9, Appli
49	37	27.2	358	4	US-08-915-795-8	Sequence 8, Appli
50	37	27.2	360	4	US-08-875-573-20	Sequence 20, Appli
51	37	27.2	360	4	US-09-232-878-2	Sequence 2, Appli
52	37	27.2	360	4	US-09-045-583-55	Sequence 55, Appli
53	37	27.2	390	2	US-08-873-093-3	Sequence 3, Appli
54	37	27.2	478	2	US-08-873-093-1	Sequence 1, Appli
55	37	27.2	478	2	US-08-873-093-4	Sequence 4, Appli
56	37	27.2	570	2	US-08-669-524-3	Sequence 3, Appli
57	37	27.2	570	2	US-08-669-524-4	Sequence 4, Appli
58	37	27.2	570	4	US-09-068-960-2	Sequence 2, Appli
59	37	27.2	570	4	US-09-068-960-4	Sequence 4, Appli
60	37	27.2	570	4	US-09-068-960-6	Sequence 6, Appli
61	37	27.2	570	4	US-09-068-960-8	Sequence 8, Appli
62	37	27.2	570	4	US-09-068-960-10	Sequence 10, Appli
63	37	27.2	643	5	PCT-US93-00031-19	Sequence 19, Appli
64	37	27.2	644	5	PCT-US93-00031-21	Sequence 21, Appli
65	37	27.2	647	4	US-09-009-490A-91	Sequence 91, Appli
66	37	27.2	647	4	US-08-482-073-5	Sequence 5, Appli
67	37	27.2	647	5	PCT-US93-00031-11	Sequence 11, Appli
68	37	27.2	647	5	PCT-US93-00031-23	Sequence 23, Appli
69	37	27.2	650	1	US-08-224-657-97	Sequence 97, Appli
70	37	27.2	650	4	US-09-354-138-97	Sequence 97, Appli
71	37	27.2	662	1	US-08-261-304-7	Sequence 7, Appli
72	37	27.2	662	1	US-08-224-657-94	Sequence 94, Appli
73	37	27.2	662	4	US-09-354-138-94	Sequence 94, Appli
74	37	27.2	671	6	5266464-2	Patent No. 5266464
75	37	27.2	694	1	US-08-164-839-4	Sequence 4, Appli
76	37	27.2	694	1	US-08-583-799-4	Sequence 4, Appli
77	37	27.2	695	1	US-08-164-839-6	Sequence 6, Appli
78	37	27.2	695	1	US-08-583-799-6	Sequence 6, Appli
79	37	27.2	735	5	PCT-US93-00031-13	Sequence 13, Appli
80	37	27.2	736	5	PCT-US93-00031-15	Sequence 15, Appli
81	37	27.2	739	4	US-08-482-073-6	Sequence 6, Appli
82	37	27.2	739	5	PCT-US93-00031-9	Sequence 9, Appli
83	37	27.2	740	5	PCT-US93-00031-17	Sequence 17, Appli
84	37	27.2	1110	1	US-08-118-441-29	Sequence 29, Appli
85	37	27.2	1110	3	US-08-338-579A-29	Sequence 29, Appli
86	37	27.2	1110	5	PCT-US94-09851-29	Sequence 29, Appli
87	37	27.2	1487	3	US-08-840-062-7	Sequence 7, Appli
88	37	27.2	2842	1	US-07-741-940-7	Sequence 7, Appli
89	37	27.2	2842	1	US-08-289-548A-7	Sequence 7, Appli
90	37	27.2	2842	1	US-08-452-654-7	Sequence 7, Appli
91	37	27.2	2843	1	US-07-741-940-2	Sequence 2, Appli
92	37	27.2	2843	1	US-08-289-548A-2	Sequence 2, Appli
93	37	27.2	2843	1	US-08-452-654-2	Sequence 2, Appli
94	37	27.2	2843	1	US-08-452-655B-2	Sequence 2, Appli
95	37	27.2	2843	1	US-08-452-655B-7	Sequence 7, Appli
96	37	27.2	2843	3	US-08-370-235A-2	Sequence 2, Appli
97	37	27.2	2843	3	US-08-450-582-2	Sequence 2, Appli
98	37	27.2	2843	3	US-08-450-582-7	Sequence 7, Appli
99	37	27.2	2973	2	US-08-821-355A-7	Sequence 7, Appli
100	37	27.2	2973	2	US-09-003-687A-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-08-976-255-11
; Sequence 11, Application US/08976255
; Patent No. 6136581
; GENERAL INFORMATION:
; APPLICANT: Jono, Keith E.
; TITLE OF INVENTION: PLOWMAN, Gregory
; TITLE OF INVENTION: KINASE GENES AND USES
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/976,255
; FILING DATE: No. 6136581ember 21, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/031,675
; FILING DATE: No. 6136581ember 22, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 229/182
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1384 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-976-255-11

Query Match 34.6%; Score 47; DB 4; Length 1384;
Best Local Similarity 64.3%; Pred. No. 41;
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 9 KENLPSLCSXQGL 22
DB 751 CCGPLPHLCSAQL 764

RESULT 2
US-09-228-986-75
; Sequence 75, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986

; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 75
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-09-228-986-75

Query Match 33.1%; Score 45; DB 4; Length 659;
Best Local Similarity 56.2%; Pred. No. 37;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 12 NLPSPCSXQGLRAEN 27
DB 49 NLPSPCSWAGVOCEGN 64

RESULT 3
US-08-997-080-43
; Sequence 43, Application US/08997080
; Patent No. 5968524
; GENERAL INFORMATION:
; APPLICANT: WATSON, JAMES D.
; APPLICANT: TAN, PAUL L.J.
; TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Spekman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,080
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 326 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-997-080-43

Query Match 31.6%; Score 43; DB 2; Length 326;
Best Local Similarity 52.9%; Pred. No. 34;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 11 ENLPSLCSXQGLRAEN 27
DB 71 ENSPALYLLDGLRAED 87


```

RESULT 4
US-08-997-362-43
; Sequence 43, Application US/08997362
; Patent No. 5985287
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiyama, Jun
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Scott, Linda
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
; TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,362
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
; FILING DATE: June 12, 1997
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347
; FILING DATE: August 29, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 326 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-997-362-43

Query Match 31.6%; Score 43; DB 2; Length 326;
Best Local Similarity 52.9%; Pred. No. 34;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 11 ENLPSLCSXQGLRAEN 27
|||:| ||||:|
Db 71 ENSPALYLLDGLRAQED 87

RESULT 5
US-08-873-970-43
; Sequence 43, Application US/08873970
; Patent No. 6001361
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiyama, Jun
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Scott, Linda
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
; TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,362
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
; FILING DATE: June 12, 1997
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347
; FILING DATE: August 29, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 326 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-997-362-43

Query Match 31.6%; Score 43; DB 2; Length 326;
Best Local Similarity 52.9%; Pred. No. 34;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 11 ENLPSLCSXQGLRAEN 27
|||:| ||||:|
Db 71 ENSPALYLLDGLRAQED 87

RESULT 6
US-09-095-855-43
; Sequence 43, Application US/09095855
; Patent No. 6160093
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Compounds and Methods for
; Treatment and Diagnosis of Mycobacterial Infections
; NUMBER OF SEQUENCES: 208
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS

```

```

; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
; TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,970
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/705,347
; FILING DATE: 29-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 326 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-873-970-43

Query Match 31.6%; Score 43; DB 3; Length 326;
Best Local Similarity 52.9%; Pred. No. 34;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 11 ENLPSLCSXQGLRAEN 27
|||:| ||||:|
Db 71 ENSPALYLLDGLRAQED 87

RESULT 6
US-09-095-855-43
; Sequence 43, Application US/09095855
; Patent No. 6160093
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Compounds and Methods for
; Treatment and Diagnosis of Mycobacterial Infections
; NUMBER OF SEQUENCES: 208
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS

```

```
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095.855
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/705,347
; FILING DATE: 29-AUG-1996
; APPLICATION NUMBER: 08/873,970
; FILING DATE: 12-JUN-1997
; APPLICATION NUMBER: 08/997,362
; FILING DATE: 23-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 326 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-095-855-43

Query Match 31.6%; Score 43; DB 4; Length 326;
Best Local Similarity 52.9%; Pred. No. 34;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 11 ENLPSLCXQGLRAEEN 27
Db 71 ENSPALYLLDGLRAQED 87

RESULT 7
US-08-705-347A-43
; Sequence 43, Application US/08705347A
; Patent No. 6284255
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiyama, Jun
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Scott, Linda
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Speckman Picard PLLC
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/705,347A
; FILING DATE: 28-AUG-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: 206.269.0565
; TELEFAX: 206.269.0563
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 326 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-705-347A-43

Query Match 31.6%; Score 43; DB 4; Length 326;
Best Local Similarity 52.9%; Pred. No. 34;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 11 ENLPSLCXQGLRAEEN 27
Db 71 ENSPALYLLDGLRAQED 87

RESULT 8
US-09-324-542-43
; Sequence 43, Application US/09324542
; Patent No. 6328978
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L.J.
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders
; FILE REFERENCE: 11000.1007c1
; CURRENT APPLICATION NUMBER: US/09/324,542
; CURRENT FILING DATE: 1999-06-02
; EARLIER APPLICATION NUMBER: US 08/997,080
; EARLIER FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 43
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
; US-09-324-542-43

Query Match 31.6%; Score 43; DB 4; Length 326;
Best Local Similarity 52.9%; Pred. No. 34;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 11 ENLPSLCXQGLRAEEN 27
Db 71 ENSPALYLLDGLRAQED 87

RESULT 9
US-08-460-529B-2
; Sequence 2, Application US/08460529B
; Patent No. 5994103
; GENERAL INFORMATION:
; APPLICANT: OLSEN, ET AL.
; TITLE OF INVENTION: Human Stanniocalcin-alpha
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
```

;; OPERATING SYSTEM: MS-DOS
;; SOFTWARE: WORD PERFECT 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/460,529B
;; FILING DATE: June 2, 1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US94/13206
;; FILING DATE: 10 NOV 1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: MULLINS, J.G.
;; REGISTRATION NUMBER: 33,073
;; REFERENCE/DOCKET NUMBER: 325800-334 (PFI43)
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 201-994-1700
;; TELEFAX: 201-994-1744
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 251 AMINO ACIDS
;; TYPE: AMINO ACID
;; STRANDEDNESS:
;; TOPOLOGY: LINEAR
;; MOLECULE TYPE: PROTEIN
;; US-08-460-529B-2

Query Match 30.9%; Score 42; DB 2; Length 251;
Best Local Similarity 46.7%; Pred. No. 36;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 4 AITXOCKENLPSLCS 18
:: ||::|| ||||
Db 192 SVQVCEQNGWSLCS 206

RESULT 10
US-08-831-132-2
; Sequence 2, Application US/08831132
; Patent No. 6008322
; GENERAL INFORMATION:
; APPLICANT: Kuestner, Rolf E.
; APPLICANT: Konklin, Darrell C.
; APPLICANT: Lok, Si
; APPLICANT: Buddle, Michele
; APPLICANT: Downey, William
; TITLE OF INVENTION: STANNIOCALCIN-2
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/831,132
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawislak, Deborah A.
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 96-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6672
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 302 amino acids

;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-831-132-2

Query Match 30.9%; Score 42; DB 3; Length 302;
Best Local Similarity 46.7%; Pred. No. 45;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 4 AITXOCKENLPSLCS 18
:: ||::|| ||||
Db 192 SVQVCEQNGWSLCS 206

RESULT 11
US-09-416-150-2
; Sequence 2, Application US/09416150
; Patent No. 6171822
; GENERAL INFORMATION:
; APPLICANT: Kuestner, Rolf E.
; APPLICANT: Konklin, Darrell C.
; APPLICANT: Lok, Si
; APPLICANT: Buddle, Michele
; APPLICANT: Downey, William
; TITLE OF INVENTION: STANNIOCALCIN-2
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc. East
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/416,150
; FILING DATE: 11-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/831,132
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawislak, Deborah A.
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 96-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6672
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 302 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
; US-09-416-150-2

Query Match 30.9%; Score 42; DB 4; Length 302;
Best Local Similarity 46.7%; Pred. No. 45;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 4 AITXOCKENLPSLCS 18
:: ||::|| ||||
Db 192 SVQVCEQNGWSLCS 206

RESULT 12

US-08-840-062-2
; Sequence 2, Application US/08840062
; Patent No. 6117977
; GENERAL INFORMATION:
; APPLICANT: LASKY, LAURENCE A.
; TITLE OF INVENTION: TYPE C LECTINS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/840,062
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1019R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1479 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-840-062-2

Query Match 30.9%; Score 42; DB 3; Length 1479;
Best Local Similarity 46.2%; Pred. No. 2.7e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 9 CKENLPSLCSXQG 21
|::|||:|
DB 495 CNOQLPSICKKAG 507

RESULT 13
US-08-840-062-4
; Sequence 4, Application US/08840062
; Patent No. 6117977
; GENERAL INFORMATION:
; APPLICANT: LASKY, LAURENCE A.
; TITLE OF INVENTION: TYPE C LECTINS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/840,062
; FILING DATE:

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1019R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1479 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-840-062-4

Query Match 30.9%; Score 42; DB 3; Length 1479;
Best Local Similarity 46.2%; Pred. No. 2.7e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 9 CKENLPSLCSXQG 21
|::|||:|
DB 496 CNOQLPSICKKAG 508

RESULT 14
US-08-716-188-1
; Sequence 1, Application US/08716188
; Patent No. 5908829
; GENERAL INFORMATION:
; APPLICANT: KELLY, RODNEY W
; TITLE OF INVENTION: USE OF MCP-1 FOR INDUCING RIPENING OF
; TITLE OF INVENTION: THE CERVIX
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/716,188
; FILING DATE: 30-SEP-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/00733
; FILING DATE: 31-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9406463.1
; FILING DATE: 31-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 117-219
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4091
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 73 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-716-188-1

Query Match 30.1%; Score 41; DB 2; Length 73;
 Best Local Similarity 50.0%; Pred. No. 13;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 9 CKENPLSLCSXQGL 22
 | | | :|| :||
 Db 26 CYRNTSSICSNEGL 39

RESULT 15

US-08-480-449-24
 ; Sequence 24, Application US/08480449
 ; Patent No. 5688927
 ; GENERAL INFORMATION:
 ; APPLICANT: Godiska, Ronald
 ; APPLICANT: Gray, Patrick W.
 ; TITLE OF INVENTION: MACROPHAGE DERIVED CHEMOKINE
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/480,449
 ; FILING DATE:
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Gass, David A.
 ; REGISTRATION NUMBER: 38,153
 ; REFERENCE/DOCKET NUMBER: 27866/32779
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312/474-6300
 ; TELEFAX: 312/474-0448
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 24:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 96 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; OTHER INFORMATION: "I-309"
 ; US-08-480-449-24

Query Match 30.1%; Score 41; DB 1; Length 96;
 Best Local Similarity 50.0%; Pred. No. 17;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 9 CKENPLSLCSXQGL 22
 | | | :|| :||
 Db 49 CYRNTSSICSNEGL 62

RESULT 16

US-08-660-542-24
 ; Sequence 24, Application US/08660542
 ; Patent No. 5932703
 ; GENERAL INFORMATION:
 ; APPLICANT: Godiska, Ronald
 ; APPLICANT: Gray, Patrick W.
 ; TITLE OF INVENTION: MACROPHAGE DERIVED CHEMOKINE AND CHEMOKINE

; TITLE OF INVENTION: ANALOGS
 ; NUMBER OF SEQUENCES: 32
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/660,542
 ; FILING DATE:
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/558,658
 ; FILING DATE: 16-NOV-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/479,620
 ; FILING DATE: 07-JUN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Gass, David A.
 ; REGISTRATION NUMBER: 38,153
 ; REFERENCE/DOCKET NUMBER: 27866/33318
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312/474-6300
 ; TELEFAX: 312/474-0448
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 24:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 96 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; OTHER INFORMATION: "I-309"
 ; FEATURE:
 ; NAME/KEY: Protein
 ; LOCATION: 1..73
 ; US-08-660-542-24

Query Match 30.1%; Score 41; DB 2; Length 96;
 Best Local Similarity 50.0%; Pred. No. 17;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 9 CKENPLSLCSXQGL 22
 | | | :|| :||
 Db 49 CYRNTSSICSNEGL 62

RESULT 17

US-08-679-493A-162
 ; Sequence 162, Application US/08679493A
 ; Patent No. 6303295
 ; GENERAL INFORMATION:
 ; APPLICANT: Taylor, Ethan W.
 ; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS
 ; FILE REFERENCE: 55-95
 ; CURRENT APPLICATION NUMBER: US/08/679,493A
 ; CURRENT FILING DATE: 1996-07-12
 ; PRIOR APPLICATION NUMBER: 60/001203
 ; PRIOR FILING DATE: 1995-07-14
 ; PRIOR APPLICATION NUMBER: 60/003,112
 ; PRIOR FILING DATE: 1995-09-01
 ; NUMBER OF SEQ ID NOS: 216
 ; SOFTWARE: PatentIn Ver. 2.0

```
; SEQ ID NO 162
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Mus sp.
US-08-679-493A-162

Query Match          30.1%; Score 41; DB 4; Length 96;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 9 CKENLPSLCSXQGL 22
   | | | | | | | |
Db 49 CYRNTSSICSNEGL 62

RESULT 18
US-08-479-603-24
; Sequence 24, Application US/08479603
; Patent No. 6320023
; GENERAL INFORMATION:
; APPLICANT: Godiska, Ronald
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: MACROPHAGE DERIVED CHEMOKINE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,603
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 27866/32780
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 96 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: "I-309"
US-08-479-603-24

Query Match          30.1%; Score 41; DB 4; Length 96;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 9 CKENLPSLCSXQGL 22
   | | | | | | | |
Db 49 CYRNTSSICSNEGL 62

RESULT 19
US-08-813-940-4

; Sequence 4, Application US/08813940
; Patent No. 5834279
; GENERAL INFORMATION:
; APPLICANT: Rubin, Harvey
; APPLICANT: Yang, Fude
; APPLICANT: Avarbock, David
; APPLICANT: Curran, Sean
; TITLE OF INVENTION: Methods of Identifying Compounds that
; TITLE OF INVENTION: Inhibit DNA Synthesis in Mycobacterium Tuberculosis and
; TITLE OF INVENTION: Composition, Reagents and Kits for Performing the Same
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5834279ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/813,940
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,271
; FILING DATE: 07-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: UPAP-0220
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 900 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-813-940-4

Query Match          30.1%; Score 41; DB 2; Length 900;
Best Local Similarity 40.0%; Pred. No. 2.2e+02;
Matches 10; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

Qy 1 TGSALT--XOCKENLPSLCSXOGLR 23
   | | | | | | | | | | | | | |
Db 258 TGAISTRWARCAKATKSLCSTRGAR 282

RESULT 20
US-08-996-083-1
; Sequence 1, Application US/08996083A
; Patent No. 6124095
; GENERAL INFORMATION:
; APPLICANT: Magna, Holly
; APPLICANT: Schaffer, Paul
; APPLICANT: Lawton, Michael
; APPLICANT: Vocum, Sue
; APPLICANT: Mitchell, Peter
; APPLICANT: Hutchinson, Nancy
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN NUCLEOTIDE PYROPHOSPHOHYDROLASE-2
; FILE REFERENCE: PF-0420 US
; CURRENT APPLICATION NUMBER: US/08/996,083A
; CURRENT FILING DATE: 1997-12-22
; NUMBER OF SEQ ID NOS: 3
```

```
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1156
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte Clone No. 6124095: 1388013
; PUBLICATION INFORMATION:
US-08-996-083-1
```

```
Query Match      29.8%; Score 40.5; DB 3; Length 1156;
Best Local Similarity 50.0%; Pred. No. 3.6e+02;
Matches 9; Conservative 3; Mismatches 3; Indels 3; Gaps 1;
```

```
Qy      2 GSAITXQCKE---NLP SL 16
      1:|:| | | | |
Db      1097 GTAVTFQCREPPAGRP SL 1114
```

RESULT 21

```
US-09-429-516-1
; Sequence 1, Application US/09429516
```

```
; Patent No. 6251389
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Magna, Holly
```

```
; APPLICANT: Schaffer, Paul
```

```
; APPLICANT: Lawton, Michael
```

```
; APPLICANT: Yocum, Sue
```

```
; APPLICANT: Mitchell, Peter
```

```
; APPLICANT: Hutchinson, Nancy
```

```
; APPLICANT: Murry, Lynn E.
```

```
; TITLE OF INVENTION: HUMAN NUCLEOTIDE PYROPHOS
```

```
; NUMBER OF SEQUENCES: 3
```

```
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
```

```
; STREET: 3174 Porter Drive
```

```
; CITY: Palo Alto
```

```
; STATE: CA
```

```
; COUNTRY: USA
```

```
; ZIP: 94304
```

```
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Diskette
```

```
; COMPUTER: IBM Compatible
```

```
; OPERATING SYSTEM: DOS
```

```
; SOFTWARE: FastSEQ for Windows Version 2.0
```

```
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/09/429,516
```

```
; FILING DATE:
```

```
; CLASSIFICATION:
```

```
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: 08/996,083
```

```
; FILING DATE:
```

```
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Billings, Lucy J
```

```
; REGISTRATION NUMBER: 36,749
```

```
; REFERENCE/DOCKET NUMBER: PF-0420 US
```

```
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: 650-855-0555
```

```
; TELEFAX: 650-845-4166
```

```
; TELEX:
```

```
; INFORMATION FOR SEQ ID NO: 1:
```

```
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 1156 amino acids
```

```
; TYPE: amino acid
```

```
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
```

```
; IMMEDIATE SOURCE:
```

```
; LIBRARY: SATPF1008
```

```
; CLONE: 1388013
```

```
; US-09-429-516-1
```

```
Query Match      29.8%; Score 40.5; DB 4; Length 1156;
Best Local Similarity 50.0%; Pred. No. 3.6e+02;
Matches 9; Conservative 3; Mismatches 3; Indels 3; Gaps 1;
```

```
Qy      2 GSAITXQCKE---NLP SL 16
      1:|:| | | | |
Db      1097 GTAVTFQCREPPAGRP SL 1114
```

RESULT 22

```
US-09-429-516-3
```

```
; Sequence 3, Application US/09429516
```

```
; Patent No. 6251389
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Magna, Holly
```

```
; APPLICANT: Schaffer, Paul
```

```
; APPLICANT: Lawton, Michael
```

```
; APPLICANT: Yocum, Sue
```

```
; APPLICANT: Mitchell, Peter
```

```
; APPLICANT: Hutchinson, Nancy
```

```
; APPLICANT: Murry, Lynn E.
```

```
; TITLE OF INVENTION: HUMAN NUCLEOTIDE PYROPHOS
```

```
; NUMBER OF SEQUENCES: 3
```

```
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
```

```
; STREET: 3174 Porter Drive
```

```
; CITY: Palo Alto
```

```
; STATE: CA
```

```
; COUNTRY: USA
```

```
; ZIP: 94304
```

```
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Diskette
```

```
; COMPUTER: IBM Compatible
```

```
; OPERATING SYSTEM: DOS
```

```
; SOFTWARE: FastSEQ for Windows Version 2.0
```

```
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/09/429,516
```

```
; FILING DATE:
```

```
; CLASSIFICATION:
```

```
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: 08/996,083
```

```
; FILING DATE:
```

```
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Billings, Lucy J
```

```
; REGISTRATION NUMBER: 36,749
```

```
; REFERENCE/DOCKET NUMBER: PF-0420 US
```

```
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: 650-855-0555
```

```
; TELEFAX: 650-845-4166
```

```
; TELEX:
```

```
; INFORMATION FOR SEQ ID NO: 3:
```

```
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 1156 amino acids
```

```
; TYPE: amino acid
```

```
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
```

```
; IMMEDIATE SOURCE:
```

```
; LIBRARY: SATPF1002
```

```
; CLONE: 422069
```

```
; US-09-429-516-3
```

```
Query Match      29.8%; Score 40.5; DB 4; Length 1156;
Best Local Similarity 50.0%; Pred. No. 3.6e+02;
Matches 9; Conservative 3; Mismatches 3; Indels 3; Gaps 1;
```

```
Qy      2 GSAITXQCKE---NLP SL 16
      1:|:| | | | |
Db      1097 GTAVTFQCREPPAGRP SL 1114
```

```

RESULT 23
US-09-227-357-299
; Sequence 299, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12

```

```

; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 299
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-227-357-299

Query Match      29.4%; Score 40; DB 4; Length 75;
Best Local Similarity 33.3%; Pred. No. 19;
Matches 9; Conservative 5; Mismatches 11; Indels 2; Gaps 1;

QY 3 SAITXQCKENLPSTL-CSXQGLRAEEN 27
   :|:| | | | | | | | | | | | | |
Db 43 AALSHACDSLPTLTLEGCRFTGMRRNN 69

RESULT 24
US-08-256-568B-92
; Sequence 92, Application US/08256568B
; Patent No. 5846704
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT; STUYVER, LIEVEN;
; APPLICANT: ROSSAU, RUDI; VAN HEUVERSWYN, HUGO
; TITLE OF INVENTION: PROCESS FOR TYPING OF HCV
; TITLE OF INVENTION: ISOLATES
; NUMBER OF SEQUENCES: 97
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIERMAN & MUSERLIAN
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,568B
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP93/03325
; FILING DATE: 26-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP/93/402,129.6
; FILING DATE: 31-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP/92/403,222.0
; FILING DATE: 27-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: CHARLES A. MUSERLIAN
; REGISTRATION NUMBER: 19,683
; REFERENCE/DOCKET NUMBER: 410,004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 661-8000
; TELEFAX: (212) 661-8002
; INFORMATION FOR SEQ ID NO: 92:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 113 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: be95

```


Query Match 29.48; Score 40; DB 4; Length 113;

```

RESULT 27
US-08-952-127-12
; Sequence 12, Application US/08952127
; Patent No. 6211336
; GENERAL INFORMATION:
; APPLICANT: Shiloh, Yosef
; APPLICANT: Taglie, Danilo A.
; APPLICANT: Collins, Francis S.
; TITLE OF INVENTION: ATAXIA-TELANGIECTASIA GENE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 6211336thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/952.127
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,995
; REFERENCE/DOCKET NUMBER: 2290.00029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 810-539-5050
; TELEFAX: 810-539-5055
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3066 amino acids

```

```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
US-08-952-127-12

```

Query Match 29.4%; Score 40; DB 4; Length 3066;
Best Local Similarity 52.9%; Pred. No. 1.3e+03;
Matches 9: Conservative 2; Mismatches 2; Indels

```

1  RESULT      28
2  US-8406-062-6
3  ; Sequence 6, Application US/08840062
4  ; Patent No. 6117977
5  ; GENERAL INFORMATION:
6  ; APPLICANT: LASKY, LAURENCE A.
7  ; APPLICANT: WU, KAI
8  ; TITLE OF INVENTION: TYPE C LECTINS
9  ; NUMBER OF SEQUENCES: 15
10 ; CORRESPONDENCE ADDRESS:
11 ; ADDRESSEE: Genentech, Inc.
12 ; STREET: 460 Point San Bruno Blvd
13 ; CITY: South San Francisco
14 ; STATE: California
15 ; COUNTRY: USA
16 ; ZIP: 94080
17 ; COMPUTER READABLE FORM:
18 ; MEDIUM TYPE: 3.5 inch, 1.44 Mb f
19 ; COMPUTER: IBM PC compatible
20 ; OPERATING SYSTEM: PC-DOS/MS-DOS
21 ; SOFTWARE: WinPath (Genentech)
22 ; CURRENT APPLICATION DATA:
23 ; APPLICATION NUMBER: US/08/840.06

```

Query Match	29.0%;	Score 39.5;	DB 3;	Length 1449;
Best Local Similarity	33.3%;	Pred. No. 6.7e+02;		
Matches	8:	Conservative	4:	Mismatches 11:
				Indels 1:
				Gaps 1:

RESULT 29
US-09-452-817-1
; Sequence 1, Application US/09452817
; Patent No. 6342374
; GENERAL INFORMATION:

Query Match	28.7%	Score 39;	DB 4;	Length 184;
Best Local Similarity	41.2%	Pred. No. 76;		
Matches	7: Conservative	3: Mismatches	7: Indels	0: Gaps

RESULT 30
US-08-588-163-5
Sequence 5, Application US/08588163
Patent No. 5643752
GENERAL INFORMATION:
APPLICANT: Hawkins, Phillip R.
APPLICANT: Murry, Lynn E.
TITLE OF INVENTION: A NOVEL TISSUE INHIBITOR OF
METALLOPROTEINASES
TITLE OF INVENTION: METALLOPROTEINASES
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,163
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Luther, Barbara J.

; APPLICANT: Kimura, Koutarou
; APPLICANT: Patterson, Garth
; APPLICANT: Ogg, Scott
; APPLICANT: Paradis, Suzanne
; APPLICANT: Tissenbaum, Heidi
; APPLICANT: Morris, Jason
; APPLICANT: Koweek, Allison
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; FILE REFERENCE: 00786/351001
; CURRENT APPLICATION NUMBER: US/08/857,076C
; CURRENT FILING DATE: 1997-05-15
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 104
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-857-076-104

Query Match 28.7%; Score 39; DB 4; Length 370;
Best Local Similarity 35.3%; Pred. No. 1.7e+02;
Matches 6; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 9 CKENLPSLCSXQGLRAE 25
|::|::| | | |
Db 75 CQKVCPTICKSHGCTAE 91

RESULT 34
US-09-163-444-2
; Sequence 2, Application US/09163444
; Patent No. 6251596
; GENERAL INFORMATION:
; APPLICANT: Cook, James W.
; APPLICANT: Bulawa, Christine
; TITLE OF INVENTION: ASPERGILLUS N-MYRISTOYL TRANSFERASE GENES AND POLYPEPTIDES
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 07334/089001
; CURRENT APPLICATION NUMBER: US/09/163,444
; CURRENT FILING DATE: 1998-09-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus N-myrystoyl
US-09-163-444-2

Query Match 28.7%; Score 39; DB 4; Length 463;
Best Local Similarity 52.9%; Pred. No. 2.2e+02;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 10 KENLPSLCSXQGLRAE 26
|::|::| | | | |
Db 264 KNHLPSTTTPGLRPM 280

RESULT 35
US-08-688-988-42
; Sequence 42, Application US/08688988B
; Patent No. 6096545
; GENERAL INFORMATION:
; APPLICANT: Lefebvre, Daniel D.
; APPLICANT: Malboobi, Mohammad A.
; TITLE OF INVENTION: PHOSPHATE STARVATION-INDUCIBLE PROTEINS
; FILE REFERENCE: PPL96-03
; CURRENT APPLICATION NUMBER: US/08/688,988B
; CURRENT FILING DATE: 1996-07-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 42
; LENGTH: 539
; TYPE: PRT
; ORGANISM: Arabidopsis Thaliana
US-08-688-988-42

Query Match 28.7%; Score 39; DB 3; Length 539;
Best Local Similarity 60.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 9 CKENLPSLCS 18
|::|::| | | |
Db 25 CEENVPTCS 34

RESULT 36
US-08-737-715-2
; Sequence 2, Application US/08737715
; Patent No. 5958685
; GENERAL INFORMATION:
; APPLICANT: Ebina, Yousuke
; TITLE OF INVENTION: MUTANT HUMAN INSULIN RECEPTOR DNA
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/737,715
; FILING DATE: 12-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 134827/1995
; FILING DATE: 12-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Nakamura, Dean H.
; REGISTRATION NUMBER: 33,981
; REFERENCE/DOCKET NUMBER: O-43323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-7860
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1382 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-737-715-2

Query Match 28.7%; Score 39; DB 2; Length 1382;
Best Local Similarity 35.3%; Pred. No. 7.6e+02;
Matches 6; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 9 CKENLPSLCSXQGLRAE 25
|::|::| | | |
Db 215 CQKVCPTICKSHGCTAE 231

RESULT 37
US-08-683-908-1
; Sequence 1, Application US/08683908
; Patent No. 5830693
; GENERAL INFORMATION:

APPLICANT: Shimizu, Sakayu
APPLICANT: Kobayashi, Michihiko
TITLE OF INVENTION: A Regulatory Factor Involved in
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Steinberg, Raskin & Davidson, P.C.
STREET: 1140 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS-DOS Editor
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,908
FILING DATE: 19-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 185626/1995
FILING DATE: 21-JUL-1995
ATTORNEY/AGENT INFORMATION:
NAME: Davidson, Clifford M
REGISTRATION NUMBER: 32,728
REFERENCE/DOCKET NUMBER: 3821004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 768-3800
TELEFAX: (212) 382-2124
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 319 AMINO ACID RESIDUES
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
ORIGINAL SOURCE:
ORGANISM: Rhodococcus rhodochrous
INDIVIDUAL ISOLATE: J1
CELL TYPE: unicellular organism
US-08-683-908-1

Query Match 28.3%; Score 38.5; DB 2; Length 319;
Best Local Similarity 42.9%; Pred. No. 1.7e+02;
Matches 9; Conservative 6; Mismatches 3; Indels 3; Gaps 2;

Qy 10 KENL--PSLCSXQGLRAEENI 28
Db 216 RONLADPNLCASQ-IAAEHNV 235

RESULT 38
US-08-220-603A-10
; Sequence 10, Application US/08220603A
; Patent No. 5612190
; GENERAL INFORMATION:
; APPLICANT: H. Arita et al.
; TITLE OF INVENTION: A Bovine Group I Phospholipase A2 Receptor
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB format
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS ver. 2.11
; SOFTWARE: Word Perfect (ASCII file)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08220603A
; FILING DATE: 30-MAR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-72323
; FILING DATE: 30-March-1993
; APPLICATION NUMBER: JP 5-206686
; FILING DATE: 20-August-1993
; APPLICATION NUMBER: JP 6-40177
; FILING DATE: 10-March-1994
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1463
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-220-603A-11

Query Match 27.9%; Score 38; DB 1; Length 1463;
Best Local Similarity 35.3%; Pred. No. 1.2e+03;
Matches 6; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

APPLICATION NUMBER: US/08/220,603A
FILING DATE: 30-MAR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-72323
FILING DATE: 30-March-1993
APPLICATION NUMBER: JP 5-206686
FILING DATE: 20-August-1993
APPLICATION NUMBER: JP 6-40177
FILING DATE: 10-March-1994
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 505 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-220-603A-10

Query Match 27.9%; Score 38; DB 1; Length 505;
Best Local Similarity 35.3%; Pred. No. 3.5e+02;
Matches 6; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 1 TGSATYXOCKENLPSLC 17
Db 466 TGLWASEECSISMPIC 482

RESULT 39
US-08-220-603A-11
; Sequence 11, Application US/08220603A
; Patent No. 5612190
; GENERAL INFORMATION:
; APPLICANT: H. Arita et al.
; TITLE OF INVENTION: A Bovine Group I Phospholipase A2 Receptor
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB format
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS ver. 2.11
; SOFTWARE: Word Perfect (ASCII file)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/220,603A
; FILING DATE: 30-MAR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-72323
; FILING DATE: 30-March-1993
; APPLICATION NUMBER: JP 5-206686
; FILING DATE: 20-August-1993
; APPLICATION NUMBER: JP 6-40177
; FILING DATE: 10-March-1994
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1463
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-220-603A-11

Search completed: August 22, 2002, 07:50:51
Job time: 189 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 07:51:24 ; Search time 27.12 Seconds
(without alignments)
99,207 Million cell updates/sec

Title: US-09-856-199-2

Perfect score: 136

Sequence: 1 TGSATYQCKENLPSLCSXQGLRAEENI 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	130	95.6	381	1	A47327 selenoprotein P pr
2	105	77.2	385	1	OMRTSP selenoprotein P pr
3	102	75.0	380	2	T10442 selenoprotein P pr
4	53	39.0	235	2	S20000 Ig light chain pre
5	52.5	38.6	604	1	YCEC31 acetolactate synth
6	52	38.2	639	2	A32545 protein kinase C (
7	51	37.5	151	2	A82978 hypothetical prote
8	49	36.0	362	2	T42049 mom-2 protein - Ca
9	49	36.0	381	2	T34254 hypothetical prote
10	48.5	35.7	573	1	C84131 acetolactate synth
11	47	34.6	358	2	T26231 hypothetical prote
12	47	34.6	1207	2	T00378 KIAA0641 protein -
13	46	33.8	95	1	FERFNC ferredoxin [2Fe-2S
14	46	33.8	146	2	T32375 hypothetical prote
15	46	33.8	1002	2	G02488 hypothetical prote
16	45	33.1	314	2	G86820 ferrochelatase (EC
17	44.5	32.7	295	2	S01062 2S seed storage pr
18	44.5	32.7	2264	1	GNVFTB genome polyprotein
19	44	32.4	133	2	F97362 PPS enzyme IIAb, m
20	44	32.4	133	2	AH2580 PTS system, IIA co
21	44	32.4	230	2	A54501 sporulated oocyst
22	44	32.4	301	2	B83666 ABC transporter (A
23	44	32.4	339	2	T43199 probable nuclear R
24	44	32.4	434	2	T37846 probable ATP-depen
25	44	32.4	443	2	A82750 hypothetical prote
26	44	32.4	900	2	T32827 hypothetical prote
27	44	32.4	1505	2	JC4851 hypoxia-inducible
28	43.5	32.0	103	2	A36568 E7 protein - human
29	43	31.6	111	2	A71045 hypothetical prote

30	43	31.6	252	2	S08054	hypothetical prote
31	43	31.6	323	2	A40433	prephytoene pyroph
32	43	31.6	323	2	T06874	probable prenyl tr
33	43	31.6	418	2	S39833	hypothetical prote
34	43	31.6	481	2	S76820	hypothetical prote
35	42.5	31.2	588	2	A90639	hypothetical prote
36	42.5	31.2	602	2	A85490	hypothetical prote
37	42.5	31.2	1214	2	S28499	probable finger pr
38	42.5	31.2	1743	2	T26859	hypothetical prote
39	42	30.9	70	2	A59346	seed storage prote
40	42	30.9	95	2	T27288	hypothetical prote
41	42	30.9	207	1	A35685	metalloproteinase
42	42	30.9	298	2	H83163	probable transcrip
43	42	30.9	302	2	JEO357	stanniocalcin homo
44	42	30.9	377	1	MENZS5	matrix protein - s
45	42	30.9	566	2	T23926	hypothetical prote
46	42	30.9	626	2	H96810	probable RNA-bind
47	42	30.9	872	2	T25186	hypothetical prote
48	42	30.9	1479	2	T42710	mannose receptor,
49	41.5	30.5	418	2	A81228	D-amino acid dehyd
50	41.5	30.5	418	2	C82001	D-amino acid dehyd
51	41.5	30.5	1123	2	D36790	hypothetical prote
52	41	30.1	96	2	A37236	I-309 protein prec
53	41	30.1	149	2	T05796	hypothetical prote
54	41	30.1	183	2	S62700	photoassimilate-re
55	41	30.1	192	2	D96498	RAP2.6 [imported]
56	41	30.1	206	2	JC1235	transcription fact
57	41	30.1	265	2	T05668	pollen allergen ho
58	41	30.1	277	2	AE1814	site-specific DNA-
59	41	30.1	311	2	D86409	hypothetical prote
60	41	30.1	345	1	MXRWIT	nonstructural prot
61	41	30.1	445	2	C83314	probable two-compo
62	41	30.1	452	2	T48998	hypothetical prote
63	41	30.1	480	1	FWPU1B	11S globulin beta
64	41	30.1	527	2	S39549	thioglycosidase (E
65	41	30.1	571	2	A81028	nucleotide-binding
66	41	30.1	607	2	G84630	probable LRR recep
67	41	30.1	613	4	C40201	artifact-warning s
68	41	30.1	641	2	T44585	acyl-CoA oxidase h
69	41	30.1	1102	2	C64477	glutamine--fructos
70	40.5	29.8	229	2	AH0814	probable ethanolam
71	40.5	29.8	323	2	S38887	2S albumin - commo
72	40.5	29.8	450	2	S46117	probable regulator
73	40.5	29.8	640	2	T19346	hypothetical prote
74	40.5	29.8	661	2	T42754	hypothetical prote
75	40.5	29.8	1280	2	T51500	hypothetical prote
76	40	29.4	145	1	A35932	angiogenin precurs
77	40	29.4	148	2	S06019	isotocin 2 / neuro
78	40	29.4	217	2	T35654	hypothetical prote
79	40	29.4	230	2	T12872	hypothetical prote
80	40	29.4	230	2	A13470	two component resp
81	40	29.4	261	2	A71192	probable glucose 1
82	40	29.4	265	2	T47537	hypothetical prote
83	40	29.4	331	2	A03553	general l-amino ac
84	40	29.4	336	2	S64829	hypothetical prote
85	40	29.4	360	2	T08673	hypothetical prote
86	40	29.4	389	2	T04816	myb-related protei
87	40	29.4	428	2	I51680	xwt-8b - African
88	40	29.4	429	2	T14524	S-locus-specific g
89	40	29.4	460	2	T17011	polygalacturonase
90	40	29.4	490	2	T01545	hypothetical prote
91	40	29.4	533	2	T00742	ubiquitin-binding
92	40	29.4	544	2	S39550	thioglycosidase (E
93	40	29.4	544	2	S19149	thioglycosidase (E
94	40	29.4	548	2	S26149	thioglycosidase (E
95	40	29.4	603	2	A75373	probable N-acetyl
96	40	29.4	650	2	E81190	hypothetical prote
97	40	29.4	672	1	KIHUCA	protein kinase C (
98	40	29.4	672	1	KIRTC	protein kinase C (
99	40	29.4	672	1	KIMSCA	protein kinase C (
100	40	29.4	672	1	KIRBC	protein kinase C (

ALIGNMENTS

RESULT 1

A47327
selenoprotein P precursor [validated] - human
C:Species: Homo sapiens (man)
C:Date: 21-Sep-1993 #sequence_revision 01-Dec-1995 #text_change 15-Sep-2000
C:Accession: A47327; S42752
R:Hall, K.E.; Lloyd, R.S.; Burk, R.F.
Proc. Natl. Acad. Sci. U.S.A. 90, 537-541, 1993
A:Title: Conserved nucleotide sequences in the open reading frame and 3' untranslated region of the human selenoprotein P gene.
A:Reference number: A47327; MUID:93133823
A:Accession: A47327
A:Molecule type: mRNA
A:Residues: 1-381 <HIL>
A:Cross-references: GB:2111793; NID:g36425; PIDN:CAA77836.1; PID:g2654365
A:Experimental source: heart and liver
A:Note: in Genbank entry HSELP, release 117.0, PIDN:CAA77836.1, the selenocysteine UGA is encoded by the sequence UGG.
R:Akesson, B.; Bellew, T.; Burk, R.F.
Biochim. Biophys. Acta 1204, 243-249, 1994
A:Title: Purification of selenoprotein P from human plasma.
A:Reference number: S42752; MUID:94191007
A:Accession: S42752
A:Molecule type: protein
A:Residues: 20-27, 'X', 29-33 <AKE>
A:Note: mature forms of 55K and 61K were detected in plasma; the protein was shown to co-sediment with selenocysteine.
C:Genetics:
A:Gene: GDB:SEPP1; SLNP
A:Cross-references: GDB:138278; OMIM:601484
A:Map position: 5q31-5q31
C:Function:
A:Description: may act as a free-radical scavenger
C:Superfamily: selenoprotein P
C:Keywords: extracellular protein; glycoprotein; heparin binding; liver; plasma; selenocysteine
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-381/Product: selenoprotein P #status experimental
F:46,83,119,128,338/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:59,300,318,330,345,352,367,369,376,378/Modified site: selenocysteine #status predicted

Query Match 95.68; Score 130; DB 1; Length 381;
Best Local Similarity 92.99; Pred. No. 8.8e-12;
Matches 26; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGSATTXCKENLPSLCXQGLRAEENI 28
||||| ||||| ||||| ||||| |||||
DB 312 TGSATTXCKENLPSLCXQGLRAEENI 339

RESULT 2

OMRTSP
selenoprotein P precursor [validated] - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 15-Sep-2000
C:Accession: A40380; B40380; S68322
R:Hall, K.E.; Lloyd, R.S.; Yang, J.G.; Read, R.; Burk, R.F.
J. Biol. Chem. 266, 10050-10053, 1991
A:Title: The cDNA for rat selenoprotein P contains 10 TGA codons in the open reading frame.
A:Reference number: A40380; MUID:91244760
A:Accession: A40380
A:Molecule type: mRNA
A:Residues: 1-385 <HIL>
A:Cross-references: GB:M63574; NID:g206893; PIDN:AAA42129.1; PID:g206894
A:Accession: B40380
A:Molecule type: protein
A:Residues: 20-41:267-287; 316-327 <HI2>
R:Chittum, H.S.; Himeno, S.; Hall, K.E.; Burk, R.F.
Arch. Biochem. Biophys. 325, 124-128, 1996
A:Title: Multiple forms of selenoprotein P in rat plasma.
A:Reference number: S68322; MUID:96140605
A:Accession: S68322
A:Molecule type: protein

Query Match 75.0%; Score 102; DB 2; Length 380;
Best Local Similarity 78.6%; Pred. No. 1.4e-07;
Matches 22; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 TGSATTXCKENLPSLCXQGLRAEENI 28
||||| ||||| ||||| ||||| |||||
DB 312 TGSATTXCKENLPSLCXQGLRAEENI 339

RESULT 4

S20000
Ig light chain precursor - rainbow trout
C:Species: Oncorhynchus mykiss (rainbow trout)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 21-Jan-2000
C:Accession: S20000; S29580
R:Daggett, A.; Bengten, E.; Pilstroem, L.
submitted to the EMBL Data Library, March 1992
A:Reference number: S20000
A:Accession: S20000
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-235 <DAG>
A:Cross-references: EMBL:X65260; NID:g64181; PID:g64182
R:Daggett, A.; Bengten, E.; Pilstroem, L.
submitted to the EMBL Data Library, September 1992
A:Description: A cluster type organisation of the loci of the immunoglobulin light chain.
A:Reference number: S29569
A:Accession: S29580
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 122-133 <DA2>
A:Cross-references: EMBL:X68522; NID:g64180; PID:g938274
C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: immunoglobulin
F:38-115/Domain: immunoglobulin homology <IMM>

Query Match 39.08; Score 53; DB 2; Length 235;
Best Local Similarity 50.04; Pred. No. 1.9;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 TGSATIXOCKENLPISLCS 18
|||:|:|:|:|:|:|
Db 37 TGDATVLRCKENPPACS 54

RESULT 5

YCEC31
acetolactate synthase (EC 4.1.3.18) III large chain - Escherichia coli

N:Alternate names: acetoxyhydroxy-acid synthase III large chain
C:Species: Escherichia coli
C:Date: 28-Aug-1985 #sequence_revision 05-Dec-1997 #text_change 05-Nov-1999
C:Accession: E64729; S14385; S40590; A01113; I41305
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617

A:Accession: E64729

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-604 <BLAT>

A:Cross-references: GB:AE000118; GB:U00096; NID:g1786262; PIDN:AAC73188.1; PID:g1786265;
A:Experimental source: strain K-12, substrain MG1655

R:Avaya, J.A.

submitted to the EMBL Data Library, January 1991

A:Reference number: S14385

A:Accession: S14385

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 31-231, 'SV', 234-235, 'V', 237-283, 'V', 285-451, 'S', 453-536, 'RG', 540-552, 'G', 597
A:Cross-references: EMBL:X55034

R:Yura, T.; Mori, H.; Nagai, H.; Nagata, T.; Ishihama, A.; Fujita, N.; Isono, K.; Mizobu
submitted to the EMBL Data Library, December 1992

A:Description: Systematic sequencing of the Escherichia coli genome: analysis of the 0-2
A:Reference number: S40591

A:Accession: S40590

A:Molecule type: DNA

A:Residues: 31-231, 'SV', 234-235, 'V', 237-283, 'V', 285-451, 'S', 453-536, 'RG', 540-552, 'G', 597
A:Cross-references: EMBL:D10483

R:Squires, C.H.; DeFellece, M.; Devereux, J.; Calvo, J.M.

Nucleic Acids Res. 11, 5299-5313, 1983

A:Title: Molecular structure of ilvIH and its evolutionary relationship to ilvG in Esche
A:Reference number: A93482; MUID:83272971

A:Accession: A01113

A:Molecule type: DNA

A:Residues: 31-231, 'SV', 234-235, 'V', 237-283, 'V', 285-451, 'S', 453-536, 'RG', 540-552, 'G', 597
A:Cross-references: GB:X01609; NID:g41794; PIDN:CAA25755.1; PID:g41795

R:Haghn, G.W.; Squires, C.H.; DeFellece, M.; Largo, C.T.; Calvo, J.M.

J. Bacteriol. 163, 186-198, 1985

A:Title: Unusual organization of the ilvIH promoter of Escherichia coli.

A:Reference number: I41305; MUID:85234358

A:Accession: I41305

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 31-38 <RES>

A:Cross-references: GB:M10738; NID:g146472; PIDN:AAA24026.1; PID:g146473

C:Comment: Isozyme III is one of the isozymes sensitive to valine inhibition.

C:Genetics:

A:Gene: ilvI

A:Map position: 2 min

C:Complex: heterodimer; composed of large chain and small chain

C:Function:

A:Description: catalyzes formation of acetoxyhydroxy acids from the corresponding alpha-ke

A:Pathway: branched-chain amino acid biosynthesis

A>Note: thiamine pyrophosphate cofactor; magnesium required

C:Superfamily: acetolactate synthase large chain; thiamin pyrophosphate-binding domain
C:Keywords: branched-chain amino acid biosynthesis; carbon-carbon lyase; flavoprotein
F:467-515/Domain: thiamin pyrophosphate-binding domain homology <TPB>
F:81/Active site: Glu #status predicted

Query Match 38.6%; Score 52.5; DB 1; Length 604;
Best Local Similarity 40.6%; Pred. No. 5.5;
Matches 13; Conservative 2; Mismatches 8; Indels 9; Gaps 1;

Qy 2 GSATIXQCKE-----NLPISCSXOGLRA 24
| | | | | : | | | | | | | | | |
Db 245 GGATAGCHQOLKETVEALNLPVVCSLMGLGA 276

RESULT 6

A32545
protein kinase C (EC 2.7.1.1) - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 18-Jun-1999
C:Accession: A32545
R:Rosenthal, A.; Rhee, L.; Yadegari, R.; Paro, R.; Ullrich, A.; Goeddel, D.V.
EMBO J. 6, 433-441, 1987
A:Title: Structure and nucleotide sequence of a Drosophila melanogaster protein kinase
A:Reference number: A32545; MUID:87218499
A:Accession: A32545
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-639 <ROS>

A:Cross-references: GB:X05076; NID:g8352; PIDN:CAA28736.1; PID:g8353; GB:Y00042
C:Genetics:
A:Gene: FlyBase:PkC53E
A:Cross-references: FlyBase:FBgn0003091
C:Superfamily: protein kinase C alpha; protein kinase C C2 region homology; protein k
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F:46-95/Domain: protein kinase C zinc-binding repeat homology <K21>
F:111-160/Domain: protein kinase C zinc-binding repeat homology <K22>
F:161-273/Domain: protein kinase C C2 region homology <K2>
F:339-599/Domain: protein kinase homology <KIN>
F:347-355/Region: protein kinase ATP-binding motif

Query Match 38.2%; Score 52; DB 2; Length 639;
Best Local Similarity 61.5%; Pred. No. 6.9;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 5 ITXQCKENLPISLC 17
: | | | | | | | | | |
Db 148 VHARCKENVPISLC 160

RESULT 7

A82978

hypothetical protein PA5359 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: A82978
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337

A:Accession: A82978

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-151 <STO>

A:Cross-references: GB:AE004947; GB:AE004091; NID:g9951665; PIDN:AAG08744.1; GSPDB:GN
A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA5359

A;Gene: KIAA0641

Query Match	34.68	Score 47;	DB 2;	Length 1207;
Best Local Similarity	64.38	Pred. No. 70;		
Matches 9; Conservative	0;	Mismatches 5;	Indels 0;	Gaps 0;

Qy 9 CKENLPSCSXQGL 22
Db 574 CCPGLPHLCSAQGL 587

RESULT 13

ferredoxin [2Fe-2S] - Rhodobacter capsulatus
 N:Alternate names: ferredoxin IV; plant-type ferredoxin
 C:Species: Rhodobacter capsulatus
 C:Date: 31-Dec-1990 #sequence.revision 27-Jan-1995 #text_change 11-Jun-1999
 C:Accession: S08393; B39857; A39519; S39897
 R:Saeki, K.; Miyatake, Y.; Young, D.A.; Marrs, B.L.; Matsubara, H.
 Nucleic Acids Res. 18, 1060, 1990
 A:Title: A plant-ferredoxin-like gene is located upstream of ferredoxin I gene (fdxN) of
 A:Reference number: S08393; MUID:90192101
 A:Accession: S08393
 A:Molecule type: DNA
 A:Residues: 1-95 <SAE>
 A:CROSS-references: EMBL:X51316; NID:g46141; PIDN:CAA35698.1; PID:g46142
 R:Saeki, K.; Sueetsugu, Y.; Tokuda, K.; Miyatake, Y.; Young, D.A.; Marrs, B.L.; Matsubara
 J. Biol. Chem. 266, 12899-12895, 1991
 A:Title: Genetic analysis of functional differences among distinct ferredoxins in Rhodospirillum rubrum
 A:Reference number: A39857; MUID:91302301

Query Match 33.8%; Score 46; DB 1; Length 95;
Best Local Similarity 37.0%; Pred. No. 9.1;
Matches 10; Conservative 2; Mismatches 15; Indels 0; Gaps 0;

QY 2 GSAITXQCKENLPSCSXQGLRAEENI 28
 | | | | : | | |
Db 31 GSGITYGCREGECCGTCTHILEGSNL 57

RESULTS

T32375
hypothetical protein K10F12.2 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*
C;date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T32375
R;Wohlmann, P.; Beck, C.
submitted to the EMBL Data Library, September 1997
A;Description: The sequence of *C. elegans* cosmid K10F12.

A;Gene: CESP:K10F12.2
A;Map position: 3
A;Introns: 46/3; 109/3

Query Match 33.8%; Score 46; DB 2; Length 146;
Best Local Similarity 42.3%; Pred. No. 14;
Matches 11; Conservative 2; Mismatches 7; Indels 6; Gaps 1;

Qy. 8 QCKENLP LSCSXQGL-----RAEEN 27
: | | | | : | | | |
Db 56 ECHEELSKLCAEAGLKGOKATRAAEN 81

RESULT 15

T02488
hypothetical protein AC2g29970 [imported] - Arabidopsis thaliana
N; Alternate names: hypothetical protein F23F1.11
C: Species: Arabidopsis thaliana (mouse-ear cress)
C: Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
C: Accession: T02488; H04702
K: Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; K
A: Description: Arabidopsis thaliana chromosome II BAC F23F1 genomic sequence.
Submitted to the EMBL Data Library, August 1998

A: Map position: 2
A: Introns: 387/2: 467/3

Query Match	33.8%	Score 46;	DB 2;	Length 1002;
Best Local Similarity	60.0%;	Pred.	No. 83;	
Matches	9;	Conservative	1;	Mismatches
			5;	Indels
				Gaps

Qy 1 TGSAITXQCKENLPS 15
: | | | | | | |
Db 437 SGSMIDDOCSEKLP 451

RESULT 16

586820

A;Map position: circular chromosome

Query Match 32.4%; Score 44; DB 2; Length 133;
Best Local Similarity 47.1%; Pred. No. 25;
Matches 8; Conservative 4; Mismatches 5; Indels

```
QY 12 NLPSLCSXQGLRAEENI 28
    ||| | | : : : : :
Db 92 NLPMLIKLAGVRSEDNM 108
```

RESULT 21

A54501

sporulated oocyst antigen TA4 precursor - *Eimeria tenella* (fragment)
 C:Species: *Eimeria tenella*
 C:Date: 09-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 07-May-1999
 C:Accession: A54501
 R:Brothers, V.M.; Kuhn, I.; Paul, L.S.; Gabe, J.D.; Andrews, W.H.; Sias, S.R.; McCannan, M.J. Biochem. Parasitol. 26, 235-248, 1988
 A:Title: Characterization of a surface antigen of *Eimeria tenella* sporozoites and synthesis of a monoclonal antibody
 A:Reference number: A54501; MUID:88261435
 A:Accession: A54501
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-230 <BRO>
 A:Cross-references: GB:M21004
 C:Keywords: disulfide bond

```
Query Match      32.4%; Score 44; DB 2; Length 230;
Best Local Similarity 40.0%; Pred. No. 41;
Matches 10; Conservative 4; Mismatches 11; Indels
```

QY	3	SAITXQCKENLPSLCSXQGLRAEEN	27
		: : : :	
Db	5	TAVTLDCKEAMNKLKRAAGLPAFED	29

RESULT 22

B83686
ABC transporter (ATP-binding protein) BH0290 [imported] - Bacillus halodurans (strain C
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: B83686
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, N.; Masui, N.; Fujii, F.; Hirai
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: B83650; MUID:20512582; PMID:11058132
A;Accession: B83686
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1301 <STO>
A;Cross-references: GB:AP001508; GB:BA000004; NID:gl0172890; PIDN:BA04009.1; GSPDB:GN00
A;Experimental source: strain C-125
C;Genetics:
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

```
Query Match      32.4%; Score 44; DB 2; Length 301;
Best Local Similarity 47.1%; Pred. No. 53;
Matches 8; Conservative 4; Mismatches 5; Indels
```

```
Qy 11 ENLPSCSXQGLRAEEN 27
    :::|:| ||| |||
Db 108 KSIPEILSKVGLRGEEN 124
```

RESULT 23

T43199
probable nuclear RNA helicase (DEAD family) - fission yeast (*Schizosaccharomyces pombe*)
C;Species: *Schizosaccharomyces pombe*

C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 17-Nov-2000
C;Accession: T431199
R;Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
DNA Res. 4, 363-369, 1997
A;Title: Identification of open reading frames in Schizosaccharomyces pombe
A;Reference number: Z17323; MUID:98162722

Query Match	32.4%	Score 44;	DB 2;	Length 339;
Best Local Similarity	52.9%	Pred. No. 60;		
Matches	9;	Conservative	1;	Mismatches
			7;	Indels
				Gaps
				0;

Qy	12	NPLSLCSXQGLRAEENI	28
		:	
Db	224	NFPSICIHGGLPOEERI	24

RESULT 2A

T37846
probable ATP-dependent RNA helicase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Nov-2000
C:Accession: T37846
R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21749
A:Accession: T37846
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-434 <MUR>
A:Cross-references: EMBL:Z99162; PIDN:CA161225.1; GSPDB:GN00066; SPDB:SPAC17G6
A:Experimental source: strain 972h-; cosmid cl7G6
C:Genetics:
A:Gene: SPDB:SPAC17G6.14c
A:Map position: 1
C:Superfamily: translation initiation factor eIF-4A

Query Match	32.4%	Score 44;	DB 2;	Length 434;
Best Local Similarity	52.9%	Pred. No. 75;		
Matches	9;	Conservative	1;	Mismatches
			7;	Indels
			0;	Gaps

Qy 12 NLP SLCSXQGLRAEENI 28
| | | : | | | |
pb 319 NFPSICIHGGLPOEERI 33

RESULT 25

A82750
hypothetical protein XF0885 [Imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: A82750
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MJID:20365717
A:Note: for a complete list of authors see reference number A59328 below

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H. submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fromm, J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigh, Chado, M.A.; Madeira, A.M.B.N.; Madeira, L.M.F.; Marino, C.L.; Marques, M.V.; Martins, H.A.; Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF0885

Query Match 32.4%; Score 44; DB 2; Length 443;
 Best Local Similarity 34.8%; Pred. No. 77;
 Matches 8; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 2 GSAITXOCKENLPSLCSXQGLRA 24

Db 114 GYKVLHCQERNPSAAEVMALRA 136

RESULT 26

T32827

hypothetical protein H19M22.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T32827

R:Wilson, R.; Wamsley, P.

submitted to the EMBL Data Library, December 1997

A:Description: The sequence of C. elegans cosmid H19M22.

A:Reference number: Z21229

A:Accession: T32827

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-900 <WIL>

A:Cross-references: EMBL:AF040648; PIDN:AAB94998.1; GSPDB:GN00021; CESP:H19M22.2

A:Experimental source: strain Bristol N2; clone H19M22

C:Genetics:

A:Gene: CESP:H19M22.2

A:Map position: 3

A:Introns: 17/2; 52/1; 75/1; 160/1; 289/3; 560/1; 599/1; 708/1; 788/1

Query Match

Best Local Similarity 32.4%; Score 44; DB 2; Length 900;

Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 11 ENLPSLCSXQGLRAEEN 27

Db 784 ENVHSLCSTPRVLAQEN 800

RESULT 27

JC4851

hypoxia-inducible factor 1 alpha homolog - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 15-Aug-1996 #sequence_revision 18-Oct-1996 #text_change 21-Jul-2000

C:Accession: JC4851

R:Nambu, J.R.; Chen, W.; Hu, S.; Crews, S.T.

Gene 172, 249-254, 1996

A:Title: The Drosophila melanogaster similar bHLH-PAS gene encodes a protein related to

A:Reference number: JC4851; MUID:96269413

A:Accession: JC4851

A:Molecule type: mRNA

A:Residues: 1-1505 <NAM>

A:Cross-references: GB:U43090; NID:g1174073; PIDN:AAC47303.1; PID:g1174074

C:Genetics:

A:Gene: slma

A:Cross-references: FlyBase:FBgn0015542
 A:Map position: 3
 F:72-125/Region: helix-loop-helix #status predicted
 F:171-433/Region: PAS domain #status predicted
 F:506-635/Region: proline-rich

Query Match 32.4%; Score 44; DB 2; Length 1505;

Best Local Similarity 42.3%; Pred. No. 2.4e+02;

Matches 11; Conservative 5; Mismatches 8; Indels 2; Gaps 1;

QY 4 AITXOCKE--NLPSLCSXQGLRAEEN 27

Db 817 SITSKSPEDSSLSLCSNLSLTQEDD 842

RESULT 28

S36568

E7 protein - human papillomavirus type 49

C:Species: human papillomavirus type 49

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999

C:Accession: S36568

R:Deilus, H.; Horrmann, B.

submitted to the EMBL Data Library, August 1993

A:Description: Primer-directed sequencing of human papillomavirus types.

A:Reference number: S36469

A:Accession: S36568

A:Molecule type: DNA

A:Residues: 1-103

A:Cross-references: EMBL:X74480; NID:g937030; PIDN:CAA52580.1; PID:g937032

C:Superfamily: papillomavirus E7 protein

C:Keywords: DNA binding; early protein; transcription regulation

Query Match

Best Local Similarity 32.0%; Score 43.5; DB 2; Length 103;

Matches 12; Conservative 3; Mismatches 12; Indels 1; Gaps 1;

QY 2 GSAITXOCKENLPSLCSXQGLRAEENI 28

Db 18 GQPIDLQCYENLTAEPAEQLEAEEL 45

RESULT 29

A71045

hypothetical protein PHI649 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000

C:Accession: A71045

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Se

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogu

DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic

A:Reference number: A71000; MUID:98344137

A:Accession: A71045

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-111 <RAW>

A:Cross-references: GB:AP000006; NID:g3236133; PIDN:BAA30761.1; PID:g3258078

A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by GenBa

C:Genetics:

A:Gene: PHI649

C:Superfamily: Pyrococcus horikoshii hypothetical protein PHI649

Query Match

Best Local Similarity 31.6%; Score 43; DB 2; Length 111;

Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 3 SAITXOCKENLPSLCSXQGLR 23

Db 13 SPLKRCQEKLPINSSSQGIK 33

```
RESULT 30
S08054
hypothetical protein 4 - maize mitochondrion plasmid S-1
C:Species: mitochondrion Zea mays (maize)
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 02-Jun-2000
C:Accession: S08054
R:Paillard, M.; Sederoff, R.R.; Levings III, C.S.
EMBO J. 4, 1125-1128, 1985
A:Title: Nucleotide sequence of the S-1 mitochondrial DNA from the S cytoplasm of maize.
A:Reference number: S07183
A:Accession: S08054
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-252 <PAI>
A:Cross-references: EMBL:X02451
C:Genetics:
A:Genome: mitochondrion
C:Superfamily: Zea mays mitochondrion 28.8K hypothetical protein
C:Keywords: mitochondrion

Query Match 31.6%; Score 43; DB 2; Length 252;
Best Local Similarity 44.0%; Pred. No. 64;
Matches 11; Conservative 2; Mismatches 8; Indels 4; Gaps 1;

Qy 2 GSATXQCKENLPSCSXQGLRAEE 26
||| ||| | | | | |
Db 195 GSA----CKEGVDSSCKEEGGCEE 215

RESULT 31
A40433
prephytoene pyrophosphatase dehydrogenase (crtE) homolog - Cyanophora paradoxa
C:Species: Cyanophora paradoxa
C>Date: 28-Feb-1992 #sequence_revision 28-Feb-1992 #text_change 24-Sep-1999
C:Accession: A40433
R:Michalowski, C.B.; Loeffelhardt, W.; Bohnert, H.J.
J. Biol. Chem. 266, 11866-11870, 1991
A:Title: An ORF323 with homology to crtE, specifying prephytoene pyrophosphate dehydroge
A:Reference number: A40433; MUID:91268060
A:Accession: A40433
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-323 <MIC>
A:Cross-references: GB:M37111; GB:M61174; NID:g336638; PIDN:AAA65472.1; PID:g336639
C:Superfamily: prenyl transferase A

Query Match 31.6%; Score 43; DB 2; Length 323;
Best Local Similarity 42.1%; Pred. No. 81;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 5 ITXQCKENLPSCSXQGLR 23
: || |||| | : |
Db 294 VAIQCKENLPSPSSKEALK 312

RESULT 32
T06874
probable prenyl transferase (EC 2.5.1.-) - Cyanophora paradoxa cyanelle
C:Species: cyanelle Cyanophora paradoxa
C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 24-Sep-1999
C:Accession: T06874
R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A.
submitted to the EMBL Data Library, July 1995
A:Description: Nucleotide sequence of the cyanelle genome from Cyanophora paradoxa.
A:Reference number: Z15840
A:Accession: T06874
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-323 <STI>
```

```
A:Cross-references: EMBL:U30821; NID:g1016083; PIDN:AAA81217.1; PID:g1016130
A:Experimental source: strain Pringsheim LB555
C:Genetics:
A:Gene: preA
A:Genome: cyanelle
C:Superfamily: prenyl transferase A
C:Keywords: cyanelle; transferase

Query Match 31.6%; Score 43; DB 2; Length 323;
Best Local Similarity 42.1%; Pred. No. 81;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 5 ITXQCKENLPSCSXQGLR 23
: || |||| | : |
Db 294 VAIQCKENLPSPSSKEALK 312

RESULT 33
S39833
hypothetical protein YBL055c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YBL0511; hypothetical protein YBL0512
C:Species: Saccharomyces cerevisiae
C>Date: 16-Dec-1993 #sequence_revision 09-Sep-1994 #text_change 29-Oct-1999
C:Accession: S39833; S45790; S37334
R:Scherens, B.; el Bakkoury, M.; Vierendeels, F.; Dubois, E.; Messenguy, F.
Yeast 9, 1355-1371, 1993
A:Title: Sequencing and functional analysis of a 32 560 bp segment on the left arm of
A:Reference number: S39824; MUID:94205266
A:Accession: S39833
A:Molecule type: DNA
A:Residues: 1-418 <SCH>
A:Cross-references: EMBL:Z23261; NID:g313733; PIDN:CAA80792.1; PID:g313743
A:Experimental source: Strain S288C
R:Dubois, E.; El Bakkoury, M.; Glansdorff, N.; Messenguy, F.; Pierard, A.; Scherens,
submitted to the Protein Sequence Database, August 1994
A:Reference number: S45782
A:Accession: S45790
A:Molecule type: DNA
A:Residues: 1-418 <DUB>
A:Cross-references: EMBL:Z35816; NID:g536083; PIDN:CAA84875.1; PID:g536084; MIPS:YBLO
A:Experimental source: strain S288C
C:Genetics:
A:Map position: 2L

Query Match 31.6%; Score 43; DB 2; Length 418;
Best Local Similarity 45.7%; Pred. No. 1e+02;
Matches 16; Conservative 3; Mismatches 6; Indels 10; Gaps 3;

Qy 1 TGSATXQCKENL-PSL-----CSXQGLRAEENI 28
||||| | | | | | | | | | | | | | | |
Db 280 TGSADLQKLLNSPFIQVNGCS---LRTEENL 311

RESULT 34
S76820
hypothetical protein sl11473 - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C:Accession: S76820
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys
s.
A:Reference number: S74322; MUID:97061201
A:Accession: S76820
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-481 <KAN>
A:Cross-references: EMBL:D90916; GB:AB001339; NID:g1653715; PIDN:BAA18732.1; PID:G101
```

A; Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

```
Query Match          31.6%; Score 43; DB 2; Length 481;
Best Local Similarity 56.2%; Pred. No. 1.2e+02;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TGSATXCKENLPSL 16
|||||  |||
Db 89 TGSAITESVNANPSI 104

RESULT 35
A0639
hypothetical protein ECs0081 [imported] - Escherichia coli (strain O157:H7, substrain R
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: A90639
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: A90639
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-588 <HAY>
A;Cross-references: GB:BA000007; PIDN:BA033504.1; PID:q13359537; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RMD 0509952
C;Genetics:
A;Gene: ECs0081
C;Superfamily: acetolactate synthase large chain; thiamin pyrophosphate-binding domain H

Query Match          31.2%; Score 42.5; DB 2; Length 588;
Best Local Similarity 37.5%; Pred. No. 1.7e+02;
Matches 12; Conservative 2; Mismatches 9; Indels 9; Gaps 1;

QY 2 GSAITXCKE-----NLPSCSXQGLRA 24
|||||  |||  |||
Db 229 GGAITAGCHQKQKETEALNLPVYSSLMGLGA 260

RESULT 36
A85490
hypothetical protein ilvI [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: A85490
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: A85490
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-602 <STO>
A;Cross-references: GB:AF005174; NID:q12512777; PIDN:AAG54381.1; GSPDB:GN00145; UWGP:Z00
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: ilvI
C;Superfamily: acetolactate synthase large chain; thiamin pyrophosphate-binding domain H

Query Match          31.2%; Score 42.5; DB 2; Length 602;
Best Local Similarity 37.5%; Pred. No. 1.7e+02;
Matches 12; Conservative 2; Mismatches 9; Indels 9; Gaps 1;

QY 2 GSAITXCKE-----NLPSCSXQGLRA 24
|||||  |||  |||
Db 243 GGAITAGCHQKQKETEALNLPVYSSLMGLGA 274
```

```
RESULT 37
S28499
probable finger protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 08-Dec-2000
C;Accession: S28499
R;Hoog, C.; Schalling, M.; Grunder-Brundell, E.; Daneholt, B.
submitted to the EMBL Data Library, June 1991
A;Description: Analysis of a murine germ cell-specific transcript that encodes a puta
A;Reference number: S28499
A;Accession: S28499
A;Molecule type: mRNA
A;Residues: 1-1214 <HOO>
A;Cross-references: EMBL:X59993; NID:g57503; PIDN:CAA42610.1; PID:g57504
A;Experimental source: strain Sprague Dawley
C;Keywords: DNA binding; zinc finger

Query Match          31.2%; Score 42.5; DB 2; Length 1214;
Best Local Similarity 37.5%; Pred. No. 3.3e+02;
Matches 9; Conservative 5; Mismatches 7; Indels 3; Gaps 1;

QY 2 GSAITXQ---CKENLPSLCSXQGL 22
|||||  |||  |||
Db 155 GSSVSKQAKSCSEVSPMCPQSV 178

RESULT 38
T26859
hypothetical protein Y43F8B.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T26859
R;Ainscough, R.
submitted to the EMBL Data Library, October 1998
A;Reference number: Z20278
A;Accession: T26859
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1743 <WIL>
A;Cross-references: EMBL:AL032623; PIDN:CAA21511.1; CESP:Y43F8B.3
A;Experimental source: clone Y43F8B
C;Genetics:
A;Gene: CESP:Y43F8B.3
A;Introns: 65/1; 92/2; 128/1; 229/1; 367/1; 422/1; 486/2; 523/1; 571/1; 628/1; 857/2;

Query Match          31.2%; Score 42.5; DB 2; Length 1743;
Best Local Similarity 39.1%; Pred. No. 4.7e+02;
Matches 9; Conservative 4; Mismatches 5; Indels 5; Gaps 1;

QY 5 ITXOCKENLPSLCSXQGLRAEEN 27
|||||  |||  |||
Db 492 LTQOC-----STCYRGLQGNEN 509

RESULT 39
A59346
seed storage protein Lec2SAL large chain [imported] - tomato (fragment)
N;Alternate names: 2S albumin large subunit
C;Species: Lycopersicon esculentum (tomato)
C;Date: 31-Dec-2001 #sequence_revision 31-Dec-2001 #text_change 31-Dec-2001
C;Accession: A59346
R;Oguri, S.
submitted to the Protein Sequence Database, September 2000
A;Reference number: A59346
A;Accession: A59346
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-70 <OGU>
A;Experimental source: cv. cherry; seed
A;Note: cross-reactive with anti-tomato lectin antiserum; one of four isoforms in see
```


C:Complex: heterodimer of large and small (see PIR:B59346) chains, disulfide linked
C:Keywords: heterodimer; pyroglutamic acid; seed; storage protein
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 30.9%; Score 42; DB 2; Length 70;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 9; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 8 QCKENLPSLCXQGLRAE 25
I I I I I I I I
Db 50 QTAENLPGLCLSPQRC 67

RESULT 40

T27288
hypothetical protein Y66A7A.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27288
R:Matthews, L.
submitted to the EMBL Data Library, October 1998
A:Reference number: Z20338
A:Accession: T27288
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-95 <WIL>
A:Cross-references: EMBL:AL032622; PIDN:CAA21498.1; GSPDB:GN00021; CESP:Y66A7A.3
C:Genetics:
A:Gene: CESP:Y66A7A.3
A:Map position: 3
A:Introns: 41/2; 62/1

Query Match 30.9%; Score 42; DB 2; Length 95;
Best Local Similarity 34.5%; Pred. No. 36;
Matches 10; Conservative 7; Mismatches 10; Indels 2; Gaps 1;

QY 1 TGSALTQCK--ENLPSLCXQGLRAEN 27
I I I I I I I I I I
Db 24 SGINISEICKFEHVPLSCAALGLATKN 52

Search completed: August 22, 2002, 07:51:28
Job time: 191 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 07:52:31 ; Search time 13.59 Seconds
(without alignments)
79.775 Million cell updates/sec

Title: US-09-856-199-2

Perfect score: 136

Sequence: 1 TGSATYXQCKENLPSCSXOGLRAEENI 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	130	95.6	381	1	SELP_HUMAN
2	110	80.9	402	1	SELP_BOVIN
3	105	77.2	385	1	SELP_RAT
4	100	73.5	380	1	SELP_MOUSE
5	52.5	38.6	574	1	ILVI_ECOLI
6	52	38.2	639	1	KPC1_DROME
7	49	36.0	362	1	MOM2_CAEEL
8	48.5	35.7	573	1	ILVI_HAEIN
9	46	33.8	95	1	FER4_RHOCA
10	45	33.1	314	1	HEM2_LACLA
11	44.5	32.7	295	1	28S5_HELAN
12	44.5	32.7	2264	1	POL1_TBRVS
13	44	32.4	253	1	TAA4_ETMTE
14	44	32.4	1507	1	SINA_DROME
15	43.5	32.0	103	1	VE7_HPV49
16	43	31.6	256	1	YMS4_MAIZE
17	43	31.6	323	1	PREA_CYPAA
18	43	31.6	418	1	YBF5_YEAST
19	43	31.6	1402	1	BIRG_MOUSE
20	43	31.6	1403	1	BIRG_MOUSE
21	42.5	31.2	1214	1	TIRG_RAT
22	42	30.9	207	1	TM1_BOVIN
23	42	30.9	302	1	STC2_HUMAN
24	42	30.9	302	1	STC2_MACNE
25	42	30.9	377	1	VNAT_SVS
26	41.5	30.5	1123	1	VG39_HSV11
27	41	30.1	96	1	SYO1_HUMAN
28	41	30.1	204	1	BTF3_MOUSE
29	41	30.1	206	1	BTF3_HUMAN
30	41	30.1	239	1	YTEA_BACSU
31	41	30.1	345	1	VP10_WTV
32	41	30.1	480	1	11SB_CUCMA
33	41	30.1	561	1	CCB2_DROME

34	41	30.1	808	1	SYFB_SYN7	P74764	synchococc
35	41	30.1	1102	1	EY20_METJA	Q58815	methanococc
36	40.5	29.8	229	1	EU20_SALTJ	Q92FV5	salmonella
37	40.5	29.8	450	1	THI2_YEAST	P38141	saccharomyc
38	40	29.4	145	1	ANG1_MOUSE	P21570	mus musculus
39	40	29.4	148	1	NEU2_CATCO	P15211	catostomus
40	40	29.4	197	1	STC2_CAVPO	P57675	cavia porce
41	40	29.4	428	1	WN8B_XENLA	P31291	xenopus lae
42	40	29.4	460	1	PGLR_MALDO	P48978	malus domes
43	40	29.4	544	1	MYR3_SINAL	P29092	sinapis alb
44	40	29.4	548	1	MYRO_BRANA	Q00326	brassica na
45	40	29.4	672	1	KPCA_BOVIN	P04409	bos taurus
46	40	29.4	672	1	KPCA_HUMAN	P17252	homo sapien
47	40	29.4	672	1	KPCA_RABIT	P10102	oryctolagus
48	40	29.4	672	1	KPCA_RAT	P05696	rattus norv
49	40	29.4	1335	1	XDH_DROME	P10351	drosophila
50	40	29.4	3066	1	ATM_MOUSE	Q62388	mus musculus
51	39.5	29.0	244	1	AG13_ARATH	Q38837	arabidopsis
52	39.5	29.0	336	1	YD48_METJA	Q58743	methanococc
53	39.5	29.0	644	1	VP4_BTIV1	P33428	bluetongue
54	39.5	29.0	644	1	VP4_BTIV13	P33429	bluetongue
55	39.5	29.0	644	1	VP4_BTIV2A	P33427	bluetongue
56	39.5	29.0	654	1	VP4_BTIV10	P07132	bluetongue
57	39.5	29.0	886	1	MCW2_XENLA	P55861	xenopus lae
58	39.5	29.0	895	1	MCW2_HUMAN	P49736	homo sapien
59	39.5	29.0	904	1	MCW2_MOUSE	P97310	mus musculus
60	39.5	29.0	907	1	MOP_DESGI	Q46509	desulfovibr
61	39	28.7	132	1	NEU2_STRCA	P21916	struthio ca
62	39	28.7	133	1	YQNG_CAEEL	Q09527	caenorhabdi
63	39	28.7	138	1	PSBU_SYN2	P74765	synchococc
64	39	28.7	207	1	TM1_HUMAN	P01033	homo sapien
65	39	28.7	208	1	Y396_RICPR	Q92dd5	rickettsia
66	39	28.7	210	1	IOVO_CHICK	P01005	gallus gall
67	39	28.7	258	1	2SS_RICCO	P01089	ricinus com
68	39	28.7	274	1	YQ59_CAEEL	Q09465	caenorhabdi
69	39	28.7	414	1	NOP3_YEAST	Q01560	saccharomyc
70	39	28.7	427	1	TRB1_AERPE	Q998t5	aeropyrum p
71	39	28.7	437	1	AFLR_ASPEL	P41765	aspergillus
72	39	28.7	444	1	AFLR_ASPEA	P43651	aspergillus
73	39	28.7	476	1	PPBH_PSEAE	P35483	pseudomonas
74	39	28.7	491	1	YD92_METJA	Q58787	methanococc
75	39	28.7	492	1	NMT_ASFFU	Q99vx3	aspergillus
76	39	28.7	525	1	G6P1_CHLTR	O843W2	chlamydia t
77	39	28.7	611	1	OM70_MOUSE	Q9czw5	mus musculus
78	39	28.7	638	1	VEL_HPV26	P36722	human papil
79	39	28.7	1141	1	MYPS_HUMAN	Q00872	homo sapien
80	39	28.7	1342	1	XDH_DROPS	P22811	drosophila
81	39	28.7	1344	1	XDH_DROSO	P91711	drosophila
82	39	28.7	1353	1	XDH_CALVI	P08793	calliphora
83	39	28.7	1372	1	INSR_MOUSE	P15208	mus musculus
84	39	28.7	1382	1	INSR_HUMAN	P06213	homo sapien
85	39	28.7	1383	1	INSR_RAT	P15127	rattus norv
86	39	28.7	1476	1	AT7A_CRIGR	P49015	cricketulus
87	39	28.7	1492	1	AT7A_RAT	P70705	rattus norv
88	39	28.7	3726	1	TRX_DROME	P20659	drosophila
89	39	28.7	3911	1	AKA9_HUMAN	Q99996	h a-kinase
90	38.5	28.3	201	1	RM12_MOUSE	Q9db15	mus musculus
91	38.5	28.3	205	1	3MCH_GLOB	Q97ey6	clostridium
92	38.5	28.3	292	1	TFSM_MOUSE	P23881	mus musculus
93	38.5	28.3	391	1	VGLG_TRTV	P33495	turkey rhin
94	38.5	28.3	476	1	HRA4_HUMAN	P8105	homo sapien
95	38.5	28.3	697	1	TRP_SCHPO	O13831	schizosacch
96	38.5	28.3	1036	1	PLD1_CRIGR	O08684	cricketulus
97	38.5	28.3	1155	1	ROM1_YEAST	P53046	saccharomyc
98	38.5	28.3	1487	1	MDS3_YEAST	P53094	saccharomyc
99	38.5	28.3	1694	1	CLH_DICDI	P25870	dicytosteli
100	38	27.9	121	1	RL19_CHLTR	O84031	chlamydia t

ALIGNMENTS

RESULT 1

```
SELP_HUMAN
ID SELP_HUMAN STANDARD; PRT; 381 AA.
AC P49908;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Selenoprotein P precursor (Sep).
GN SEPPI OR SELP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver, and Heart;
RX MEDLINE=93133823; PubMed=8421687;
RA Hill K.E., Lloyd R.S., Burk R.F.;
RT "Conserved nucleotide sequences in the open reading frame and 3'
RT untranslated region of selenoprotein P mRNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:537-541(1993).
RN [2]
RP PARTIAL SEQUENCE.
RX MEDLINE=94191007; PubMed=8142465;
RA Akesson B., Bellaw T., Burk R.F.;
RT "Purification of selenoprotein P from human plasma.";
RL Biochim. Biophys. Acta 1204:243-249(1994).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=98413836; PubMed=9735174;
RA Mostert V., Lombeck I., Abel J.;
RT "A novel method for the purification of selenoprotein P from human
RT plasma.";
RL Arch. Biochem. Biophys. 357:326-330(1998).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=20239644; PubMed=10775431;
RA Mostert V.;
RT "Selenoprotein P: properties, functions, and regulation.";
RL Arch. Biochem. Biophys. 376:433-438(2000).
RN [5]
RP REVIEW.
RX MEDLINE=95017128; PubMed=7931697;
RA Burk R.F., Hill K.E.;
RT "Selenoprotein P. A selenium-rich extracellular glycoprotein.";
RL J. Nutr. 124:1891-1897(1994).
CC -!- FUNCTION: MIGHT BE RESPONSIBLE FOR SOME OF THE EXTRACELLULAR
CC ANTIOXIDANT DEFENSE PROPERTIES OF SELENIUM OR MIGHT BE INVOLVED IN
CC THE TRANSPORT OF SELENIUM.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MADE IN THE LIVER AND HEART AND SECRETED INTO
CC THE PLASMA. IT IS ALSO FOUND IN THE KIDNEY.
CC -!- MISCELLANEOUS: THE SELENOCYSTEINES ARE ALL ENCODED BY THE OPAL
CC CODON, UGA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z11793; CAA77836.1; -.
DR GIM; 601484; -.
KW Glycoprotein; Signal; Selenium; Selenocysteine; Plasma.
FT SIGNAL 1 19
FT CHAIN 20 381 SELENOPROTEIN P.
FT DOMAIN 244 249 POLY-HIS.
FT SE_CYS 59 59
FT SE_CYS 300 300
FT SE_CYS 318 318
FT SE_CYS 330 330
FT SE_CYS 345 345
FT SE_CYS 345 345

Query Match 95.6%; Score 130; DB 1; Length 381;
Best Local Similarity 92.9%; Pred. No. 6.2e-13;
Matches 26; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGSATXOCKENLPSLC SXQGLRAEENI 28
DB 312 TGSATXOCKENLPSLC SCQGLRAEENI 339

RESULT 2
SELP_BOVIN
ID SELP_BOVIN STANDARD; PRT; 402 AA.
AC P49907; O19003;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Selenoprotein P-like protein precursor.
GN SEPPI OR SELP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RX MEDLINE=95364621; PubMed=7637580;
RA Saijoh K., Saito N., Lee M.J., Fujii M., Kobayashi T., Sumino K.;
RT "Molecular cloning of cDNA encoding a bovine selenoprotein P-like
RT protein containing 12 selenocysteines and a (His-Pro) rich domain
RT insertion, and its regional expression.";
RL Brain Res. Mol. Brain Res. 30:301-311(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98019090; PubMed=9358058;
RA Fujii M., Saijoh K., Kobayashi T., Fujii S., Lee M.J., Sumino K.;
RT "Analysis of bovine selenoprotein P-like protein gene and availability
RT of metal responsive element (MRE) located in its promoter.";
RL Gene 199:211-217(1997).
CC -!- FUNCTION: IT CONSTITUTES A MAJOR SELENIUM POOL IN THE BRAIN AND
CC MAY PLAY AN IMPORTANT ROLE IN DEVELOPING AND/OR MODULATING THE
CC MORPHOLOGY OF NEURONS AND/OR GLIAL CELLS.
CC -!- SUBCELLULAR LOCATION: Extracellular (By similarity).
CC -!- TISSUE SPECIFICITY: BRAIN AND KIDNEY. MOST PROMINENTLY EXPRESSED
CC IN THE CEREBELLAR CORTEX, HIPPOCAMPUS AND OLFACTORY BULB.
CC -!- MISCELLANEOUS: THE SELENOCYSTEINES ARE ALL ENCODED BY THE OPAL
CC CODON, UGA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D25220; BAA04949.2; -.
DR EMBL; D8033; BAA23414.1; -.
DR EMBL; D88031; BAA23414.1; JOINED.
```



```

CC -!- TISSUE SPECIFICITY: MADE IN THE LIVER AND SECRETED INTO THE
CC PLASMA (BY SIMILARITY).
CC -!- MISCELLANEOUS: THE SELENOCYSTEINES ARE ALL ENCODED BY THE OPAL
CC CODON, UGA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X99807; CAA68140.1; -.
CC DR MGD; MG1:894288; Sep1.
CC KW Glycoprotein; Signal; Selenium; Selenocysteine; Plasma.
CC FT SIGNAL 1 19 BY SIMILARITY.
CC FT CHAIN 20 380 SELENOPROTEIN P.
CC FT SE_CYS 59 59
CC FT SE_CYS 259 259
CC FT SE_CYS 277 277
CC FT SE_CYS 318 318
CC FT SE_CYS 330 330
CC FT SE_CYS 352 352
CC FT SE_CYS 366 366
CC FT SE_CYS 368 368
CC FT SE_CYS 375 375
CC FT SE_CYS 377 377
CC FT CARBOHYD 83 83 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 365 365 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 370 370 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SQ SEQUENCE 380 AA; 42236 MW; 72F7031941FA7212 CRC64;

Query Match 73.5%; Score 100; DB 1; Length 380;
Best Local Similarity 71.4%; Pred. No. 2.9e-08;
Matches 20; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 TGSATTXCKENLPSCSXQGLRAEENI 28
Db 312 SGSAIACQCAENLPSCSCQGLFAEEKV 339

RESULT 5
ID ILVI_ECOLI STANDARD; PRT; 574 AA.
AC P00893; P78045;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Acetolactate synthase isozyme III large subunit (EC 4.1.3.18) (AHAS-
DE III) (Acetohydroxy-acid synthase III large subunit) (ALS-III).
DE ILVI OR B0077.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83272971; PubMed=6308579;
RA Squires C.H., Defelice M., Devereux J., Calvo J.M.;
RT "Molecular structure of ilvIH and its evolutionary relationship to
RT ilvG in Escherichia coli K12."
RL Nucleic Acids Res. 11:5299-5313(1983).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

```

```

RA Mau B., Shao Y.;
RL "The complete genome sequence of Escherichia coli K-12.";
RN Science 277:1453-1474(1997).
RP [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=92334977; PubMed=1630901;
RY Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
RY Isono K., Mizobuchi K., Nakata A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of
RT the 0-2.4 min region.";
RL Nucleic Acids Res. 20:3305-3308(1992).
RN [4]
RP SEQUENCE FROM N.A.
RA Ayala J.A.;
RT "Regulation of transcription at the 2-minute region of the genetic map
RT of Escherichia coli.";
RL Submitted (JAN-1991) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-8 FROM N.A.
RX MEDLINE=85234358; PubMed=3891724;
RY Haughn G.W., Squires C.H., Defelice M., Largo C.T., Calvo J.M.;
RT "Unusual organization of the ilvIH promoter of Escherichia coli.";
RL J. Bacteriol. 163:186-198(1985).
RN [6]
RP SEQUENCE OF 1-12.
RC STRAIN=K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RY Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RT in the genome of Escherichia coli K-12.";
RL Electrophoresis 18:1259-1313(1997).
CC -!- CATALYTIC ACTIVITY: 2-ACETOLACTATE + CO(2) = 2 PYRUVATE (THIS
CC ENZYME ALSO CATALYZES FORMATION OF 2-ACETO-2-HYDROXYBUTANOATE).
CC -!- COFACTOR: THIAMINE PYROPHOSPHATE, AND MAGNESIUM.
CC -!- ENZYME REGULATION: SENSITIVE TO VALINE INHIBITION.
CC -!- PATHWAY: FIRST STEP IN VALINE AND ISOLEUCINE BIOSYNTHESIS.
CC -!- SUBUNIT: DIMER OF LARGE AND SMALL CHAINS.
CC -!- MISCELLANEOUS: E. COLI CONTAINS GENES FOR 3 AHAS ISOZYMES: ILVEN,
CC ILVG AND ILVIH.
CC -!- SIMILARITY: WITH OTHER ENZYMES WHICH REQUIRE TPP.
CC -!- CAUTION: REF.1 AND 3 SEQUENCES DIFFER FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 523.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X01609; CAA25755.1; ALT_FRAME.
CC EMBL; AE000118; AAC73188.1; ALT_INIT.
CC EMBL; D10483; BAA01345.1; ALT_FRAME.
CC EMBL; X55034; CAA38854.1; -.
CC EMBL; M10738; AAA24026.1; -.
CC PIR; A01113; YCEC31.
CC PIR; S14385; S14385.
CC EcoGene; EG10500; ilvI.
CC InterPro; IPR000399; TPP_enzyme.
CC Pfam; PF00205; TPP_enzymes_1.
CC Pfam; PF02775; TPP_enzymes_C; 1.
CC Pfam; PF02776; TPP_enzymes_N; 1.
CC PROSITE; PS00187; TPP_ENZYMES; 1.
CC Branched-chain amino acid biosynthesis; Flavoprotein; Magnesium;
CC Thiamine pyrophosphate; Lyase; Complete proteome.
CC ACT_SITE 51 51 BY SIMILARITY.
CC CONFLICT 7 7 A -> G (IN REF. 3).
CC CONFLICT 202 203 TL -> SV (IN REF. 1 AND 3).
CC CONFLICT 206 206 A -> V (IN REF. 1 AND 3).
CC CONFLICT 254 254 A -> V (IN REF. 1 AND 3).
CC CONFLICT 422 422 T -> S (IN REF. 1 AND 3).

```

```

RESULT      7
MOM2_CAEL
ID      MOM2_CAEL      STANDARD;      PRT;      362 AA.
AC      Q10459; O16146;
DT      16-OCT-2001 (Rel. 40, Created)
DT      15-OCT-2001 (Rel. 40, Last sequence update)
DT      01-MAR-2002 (Rel. 41, Last annotation update)
DE      Mom-2 protein precursor.
DE      MOM-2 OR F38E1.7.
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC      Rhabditidae; Peloderinae; Caenorhabditis.
OX      NCBI_Taxid=6239;
RN      [1]
RP      SEQUENCE FROM N.A.
RP      STRAIN=BRISTOL N2;
RC      MEDLINE=97433081; PubMed=9288750;
RA      Rocheleau C.E., Downs W.D., Lin R., Wittmann C., Bei Y., Cha Y.-H.;
RT      Ali M., Pries J.R., Mello C.C.;
RT      "Wnt signaling and an APC-related gene specify endoderm in early C.
RT      elegans embryos.";
RT      Cell 90:707-716(1997).
RN      [2]
RP      SEQUENCE FROM N.A.
RP      STRAIN=BRISTOL N2;
RC      Gattung S., Le T.;
RA      Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      REVISIONS.
RA      Waterston R.;
RA      Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC      -!- FUNCTION: LIGAND FOR MEMBERS OF THE FRIZZLED FAMILY OF SEVEN
CC      TRANSMEMBRANE RECEPTORS. PROBABLE DEVELOPMENTAL PROTEIN INVOLVED
CC      IN ENDODERM SPECIFICATION AND CLEAVAGE AXIS DETERMINATION.

```



```

CC CC -!- COFACTOR: BINDS 1 2FE-2S CLUSTER (PROBABLE).
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC DR EMBL; X51316; CAA35698.1; -
CC DR EMBL; M59855; AAA26109.1; -
CC DR EMBL; X72888; CAA51403.1; -
CC DR PIR; S08393; S08393.
CC DR PIR; A39519; A39519.
CC DR HSP; P00256; 2CJO.
CC DR InterPro; IPR000564; 2Fe2s_ferredoxin.
CC DR InterPro; IPR001041; Ferredoxin.
CC DR Pfam; PF00111; fer2; 1.
CC DR PROSITE; PS00197; 2Fe2s_FERREDOXIN; 1.
CC KW Electron transport; Iron-sulfur; Nitrogen fixation.
CC FT METAL 38 38 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC FT METAL 43 43 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC FT METAL 46 46 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC FT METAL 81 81 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC SQ SEQUENCE 95 AA; 10163 MW; 7CD077A779D543F0 CRC64;

Query Match 33.8%; Score 46; DB 1; Length 95;
Best Local Similarity 37.0%; Pred. No. 1.9;
Matches 10; Conservative 2; Mismatches 15; Indels 0; Gaps 0;

Qy 2 GSAITXCKENLPISLCSXQGLRAEENI 28
  ||| ||| | | | | | | | | | |
Db 31 GSGITYGCGEGCGTCWTHILGSENL 57

RESULT 10
HEMZ_LACLA
ID HEMZ_LACLA STANDARD; PRT; 314 AA.
AC Q9CFB4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ferredoxinase (EC 4.99.1.1) (Protoheme ferro-lyase) (Heme
DE synthetase).
GN HEHM OR LL1567.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RA MEDLINE=21235186; PubMed=1137471;
RA Bolotin A., Wincker P., Mauer S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RA "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403";
RL Genome Res. 11:731-753(2001)
CC -!- FUNCTION: CATALYZES THE FERROUS INSERTION INTO PROTOPORPHYRIN IX.
CC CC -!- CATALYTIC ACTIVITY: Protoporphyrin + Fe(2+) = protoheme + 2 H(+).
CC CC -!- PATHWAY: LAST STEP IN PROTOHEME BIOSYNTHESIS.
CC CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC CC -!- SIMILARITY: BELONGS TO THE FERROCHELATASE FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).

```

```

CC CC EMBL; AE006387; AAK05665.1; -
CC DR InterPro; IPR001015; Ferrochelatase.
CC DR Pfam; PF00762; Ferrochelatase; 1.
CC DR PROSITE; PS00534; FERROCHELATASE; FALSE_NEG.
CC KW Porphyrin biosynthesis; Heme biosynthesis; Lyase; Iron;
CC CC complete proteome.
CC FT METAL 186 186 IRON (BY SIMILARITY).
CC FT METAL 268 268 IRON (BY SIMILARITY).
CC SQ SEQUENCE 314 AA; 36436 MW; 00B03F49869B5E1F CRC64;

Query Match 33.1%; Score 45; DB 1; Length 314;
Best Local Similarity 47.8%; Pred. No. 9.2;
Matches 11; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

Qy 4 ATTXXCKENLPISLCSXQGLRAEE 26
  ||||| : |||||
Db 199 AYECKETTRLVLSKGLREE 221

RESULT 11
2SS5_HELAN
ID 2SS5_HELAN STANDARD; PRT; 295 AA.
AC P15461;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE 2S seed storage protein precursor (2S albumin storage protein).
GN HAG5.
OS Helianthus annuus (Common sunflower).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
OC Heliantheae; Helianthus.
OX NCBI_TaxID=4232;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 162-173.
RC STRAIN=CV. GIANT GREY STRIPE;
RX MEDLINE=88142538; PubMed=2830455;
RA Allen R.D., Cohen E.A., Vonder Haar R.A., Adams C.A., Ma D.P.,
RA Nessler C.L., Thomas T.L.;
RA "Sequence and expression of a gene encoding an albumin storage
RT protein in sunflower.";
RL Mol. Gen. Genet. 210:211-218(1987).
CC -!- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.
CC CC -!- PTM: IT IS POSSIBLE THAT THE 38 kDa PRECURSOR IS CLEAVED INTO TWO
CC CC POLYPEPTIDES THAT ARE APPROXIMATELY THE SAME SIZE. THE MATURE
CC CC PROTEIN IS COMPOSED OF A SINGLE POLYPEPTIDE CONTAINING ONE OR MORE
CC CC INTRA-MOLECULAR DISULFIDE LINKAGES.
CC CC -!- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC DR EMBL; X06410; CAA29699.1; -
CC DR PIR; S01062; S01062.
CC DR InterPro; IPR003612; AA1.
CC DR InterPro; IPR001768; Cereal_ryp_amy1_inh.
CC DR Pfam; PF00234; ryp_alpha_amy1; 2.
CC DR SMART; SM00499; AA1; 2.
CC KW Seed storage protein; Albumin; Signal; Multigene family.
CC FT SIGNAL 1 20 POTENTIAL.
CC FT PROPEP 21 161 POTENTIAL.
CC FT CHAIN 162 295 2S SEED STORAGE PROTEIN.
CC SQ SEQUENCE 295 AA; 34071 MW; 8958A106805142A1 CRC64;

```

Query Match 32.7%; Score 44.5; DB 1; Length 295;
Best Local Similarity 40.7%; Pred. No. 10;
Matches 11; Conservative 4; Mismatches 9; Indels 3; Gaps 1;

QY 2 GSATXOCKEN--LPSLCSEXOGLRAE 25
||| :| | | | | | | | | |
Db 123 GSQTQLKQKAQILPNCVLSRRCE 149

RESULT 12
ID POLL_TBRSVS STANDARD; PRT; 2264 AA.
AC P18522;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE RNAL polypeptide (250 kDa protein) [Contains: 63 kDa protease
cofactor; 72 kDa membrane-binding protein; Genome-linked protein
(VPG); Protease (EC 3.4.22.-); RNA-directed RNA polymerase (EC
2.7.7.48)].
OS Tomato black ring virus (strain S) (TBV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Comoviridae;
OX Nepovirus.
ON NCBI_TaxID=12279;
RN [1]
RP SEQUENCE FROM N.A.
RA Greif C., Hemmer O., Fritsch C.;
RT "Nucleotide sequence of tomato black ring virus RNA-1.";
RL J. Gen. Virol. 69:1517-1529(1988).
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: WITH RNA-1 POLYPEPTIDE OF HUNGARIAN GRAPEVINE CHROME
MOZAIC VIRUS (GCMV) AND WITH POLYPEPTIDE B OF COMPEA MOSAIC VIRUS
(CPMV).

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL: D00322; BAA00234.1; --
DR PIR: JQ0009; GNVTB.
DR MEROPS: C03.004; --
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR001205; RNA_pol_P3D.
DR Pfam: PF00680; RNA_dep_RNA_pol_1.
DR Pfam: PF00910; RNA_helicase_1.
KW Polyprotein; Transmembrane; Hydrolase; Protease; Transferase;
RNA-directed RNA polymerase; ATP-binding.
FT CHAIN 1 565 63 KDA PROTEASE COFACTOR.
FT CHAIN 566 1212 72 KDA MEMBRANE-BINDING PROTEIN.
FT CHAIN 1213 1232 GENOME-LINKED PROTEIN.
FT CHAIN 1233 1440 CYSTEINE PROTEASE.
FT CHAIN 1441 2264 RNA-DIRECTED RNA POLYMERASE.
FT NP_BIND 780 787 ATP (POTENTIAL).
FT ACT_SITE 1400 1400 CYSTEINE PROTEASE (POTENTIAL).
SQ SEQUENCE 2264 AA; 253676 MW; 65949E4B5CE8B722 CRC64;

Query Match 32.7%; Score 44.5; DB 1; Length 2264;
Best Local Similarity 46.2%; Pred. No. 79;
Matches 12; Conservative 3; Mismatches 10; Indels 1; Gaps 1;

QY 1 TGSATXOCKEN-LPSLCSEXOGLRAE 25
||| :| | | | | | | | | |
Db 294 TGSNVTSRKPNNVLNLCPFMGLGAK 319

RESULT 13
ID TA4_EIMTE STANDARD; PRT; 253 AA.
TA4_EIMTE

Query Match 32.7%; Score 44; DB 1; Length 253;
Best Local Similarity 40.0%; Pred. No. 11;
Matches 10; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 3 SATXOCKENLPSLCSEXOGLRAEEN 27
:||| ||| :| | | | | | | | | |
Db 28 TAVTLDCKEAMNKLRKAAGLPAPED 52

RESULT 14
SIMA_DROME STANDARD; PRT; 1507 AA.
ID SIMA_DROME
AC Q24167; Q9VAA5;
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Similar protein.
GN SIMA OR CG7951.
OS Drosophila melanogaster (Fruit fly).

AC P13399;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Sporulated oocyst TA4 antigen precursor (Major sporozoite surface
antigen).
DE Elmeria tenella.
OS Eukaryota; Alveolata; Apicomplexa; Coccidia; Elmeriida; Elmeriidae;
OC Elmeria.
OX NCBI_TaxID=5802;
RN [1]
RP SEQUENCE FROM N.A.
RA Files J.G., Paul L.S., Gabe J.D.;
RT "Identification and characterization of the gene for a major surface
antigen of Elmeria tenella.";
RL (In) Agabian N., Goodman H., Noguera N. (eds.);
RL Molecular strategies of parasitic invasion, UCLA symposia on molecular
and cellular biology, pp.713-723, Alan R. Liss, New York (1987).
RN [2]
RP SEQUENCE OF 23-253 FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=27 / SALESBURY LAB.;
RX MEDLINE=88261435; PubMed=3290678;
RA Brothers V.M., Kuhn I., Paul L.S., Gabe J.D., Andrews W.H.,
Sias S.R., McCaman M.T., Dragon E.A., Files J.G.;
RT "Characterization of a surface antigen of Elmeria tenella sporozoites
and synthesis from a cloned cDNA in Escherichia coli.";
RL Mol. Biochem. Parasitol. 28:235-248(1988).
CC -1- SUBUNIT: THE TA4 ANTIGEN IS COMPOSED OF A 17 kDa AND A 8 kDa
CHAIN, LINKED BY A DISULFIDE BOND.
CC -1- SUBCELLULAR LOCATION: SURFACE OF SPOROZOITES.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AFTER INITIATION OF OOCYST
SPORULATION.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL: M21088; AAA29074.1; --
DR EMBL: M21004; AAA29075.1; --
DR PIR: A54501; A54501.
KW Antigen; Signal.
FT SIGNAL 1 23
FT CHAIN 24 181 SPORULATED OOCYST TA4 ANTIGEN 17 KDA
SUBUNIT.
FT PROPEP 182 184 MISSING IN PROCESSED ANTIGEN.
FT CHAIN 185 253 SPORULATED OOCYST TA4 ANTIGEN 8 KDA
SUBUNIT.
FT SEQUENCE 253 AA; 27137 MW; 6DBDFB4102F70AE CRC64;

Query Match 32.4%; Score 44; DB 1; Length 253;
Best Local Similarity 40.0%; Pred. No. 11;
Matches 10; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 3 SATXOCKENLPSLCSEXOGLRAEEN 27
:||| ||| :| | | | | | | | | |
Db 28 TAVTLDCKEAMNKLRKAAGLPAPED 52

RESULT 14
SIMA_DROME STANDARD; PRT; 1507 AA.
ID SIMA_DROME
AC Q24167; Q9VAA5;
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Similar protein.
GN SIMA OR CG7951.
OS Drosophila melanogaster (Fruit fly).

DR	ENBL; AEO03772; AAF57008.2; -.
DR	FlyBase; FBgn0015542; sima.
DR	InterPro; IPR003015; HLH_Myc.
DR	InterPro; IPR001092; HLH_dim.
DR	InterPro; IPR001610; PAC.
DR	InterPro; IPR000014; PAS.
DR	Pfam; PF00785; PAC; 1.
DR	Pfam; PF00989; PAS; 2.
DR	SMART; SM00353; HLH; 1.
DR	SMART; SM00086; PAC; 1.
DR	SMART; SM00091; PAS; 2.
DR	PROSITE; PS00038; HELIX_LOOP_HELIX; 1.
DR	PROSITE; PS0112; PAS; 2.
KW	Repeat; DNA-binding; Nuclear protein; Transcription regulation;
KW	Activator; Coiled coil.
FT	DNA_BIND 72 85 DOMAIN 86 126
FT	DOMAIN 167 240
FT	DOMAIN 307 377
FT	DOMAIN 381 422
FT	DOMAIN 577 587
FT	DOMAIN 860 908
FT	DOMAIN 982 1054
FT	DOMAIN 1110 1162
FT	DOMAIN 26 39
FT	DOMAIN 718 725
FT	DOMAIN 759 763
FT	DOMAIN 767 776
FT	DOMAIN 907 918
FT	DOMAIN 945 948
FT	DOMAIN 990 998
FT	DOMAIN 1020 1038
FT	DOMAIN 1113 1126
FT	DOMAIN 1146 1162
FT	DOMAIN 1205 1208
FT	DOMAIN 1277 1284
FT	DOMAIN 1298 1301
FT	CONFLICT 38 38
FT	CONFLICT 345 345
FT	CONFLICT 492 492
FT	CONFLICT 588 588
FT	CONFLICT 709 709
FT	CONFLICT 776 776
FT	CONFLICT 895 895
FT	CONFLICT 902 902
FT	CONFLICT 982 982
FT	CONFLICT 1125 1126
FT	CONFLICT 1154 1157
FT	CONFLICT 1444 1444
FT	CONFLICT 1447 1447
FT	CONFLICT 1451 1451
FT	CONFLICT 1494 1494
SQ	SEQUENCE 1507 AA; 165824 MW; 4102939C8FBFBC6 CRC64; S -> A (IN REF. 1). S -> L (IN REF. 1). A -> V (IN REF. 1). T -> I (IN REF. 1). T -> K (IN REF. 1). Q -> QQQQ (IN REF. 1). Q -> QQ (IN REF. 1). G -> S (IN REF. 1). A -> T (IN REF. 1). MISSING (IN REF. 1). MISSING (IN REF. 1). E -> L (IN REF. 1). G -> C (IN REF. 1). S -> N (IN REF. 1). D -> G (IN REF. 1).
Query Match	32.4%; Score 44; DB 1; Length 1507;
Best Local Similarity	42.3%; Pred.No. 63;
Matches 11; Conservative 5; Mismatches 8; Indels 2; Gaps	
QY	4 AIITXCKE--NLPSLCSXOGLRAEN 27 :: :- - l ll l l l l l : Db 814 SITSKSPEDSLPISLCSNLSLTQEDD 839
RESULT 15	
VE7_HPVA9	ID VE7_HPVA9 STANDARD; PRT; 103 AA.
IC P36830:	
DT 01-JUN-1994 (Rel. 29, Created)	
DT 01-JUN-1994 (Rel. 29, Last sequence update)	
DT 01-WAR-2002 (Rel. 41, Last annotation update)	
DE E7 protein.	
GN EF7	

```
OS Human papillomavirus type 49.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10616;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94265501; PubMed=8205838;
RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC -!- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING
CC ACTIVITIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X74480; CAA52580.1; -.
DR PIR: S36568; S36568.
DR InterPro: IPR000148; Papvi_E7.
DR Pfam: PF00527; E7; 1.
KW Early protein; Transcription regulation; Oncogene;
KW DNA-binding; Trans-acting factor.
FT DOMAIN 58 61 C-XX-C MOTIF 1.
FT DOMAIN 91 94 C-XX-C MOTIF 2.
FT SEQUENCE 103 AA; 11454 MW; 5235514B2926DDA0 CRC64;
SQ
Query Match 32.0%; Score 43.5; DB 1; Length 103;
Best Local Similarity 42.9%; Pred. No. 5.2;
Matches 12; Conservative 3; Mismatches 12; Indels 1; Gaps 1;

OY 2 GSAITXQCKENLPSC-SXQGLRAEENI 28
I | | | | | : | | | | : | | | | :
D 18 GQPIDLCQVENLTAPAPQEELEABEEL 45

RESULT 16
ID YMS4_MAIZE STANDARD; PRT; 256 AA.
AC P10580;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 29 kDa protein in mitochondrial S-1 DNA (URF 4).
OS Zea mays (Maize).
OG Mitochondrion.
OG Plasmid S-1.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Paillard M., Sederoff R.R., Levings C.S. III;
RT Nucleotide sequence of the S-1 mitochondrial DNA from the S cytoplasm
RT of maize.;
RL EMBL J. 4:1125-1128(1985).
CC -!- MISCELLANEOUS: THE MITOCHONDRIA FROM THE S MALE-STERILE CYTOPLASM
CC OF MAIZE CONTAIN UNIQUE DNA-PROTEIN COMPLEXES, DESIGNATED S-1 AND
CC S-2. THESE COMPLEXES CONSIST OF DOUBLE-STRANDED LINEAR DNAs WITH
CC PROTEINS COVALENTLY ATTACHED TO THE 5' TERMINI.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X02451; -: NOT_ANNOTATED_CDS.
DR MaizeDB: 69620; -.
KW Hypothetical protein; Mitochondrion; Plasmid.
SQ SEQUENCE 256 AA; 29338 MW; BID6DB31E402D175 CRC64;

Query Match 31.6%; Score 43; DB 1; Length 256;
Best Local Similarity 44.0%; Pred. No. 15;
Matches 11; Conservative 2; Mismatches 8; Indels 4; Gaps 1;

OY 2 GSAITXQCKENLPSCSXQGLRAEE 26
I | | | | | : | | | | : | | | | :
D 199 GSA-----CKEGVDSCKEEGGGCEE 219

RESULT 17
ID PREA_CYPAPA STANDARD; PRT; 323 AA.
AC P31171;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Prenyl transferase (EC 2.5.1.-).
OS Cyanophora paradoxa.
OG Cyanelle.
OC Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.
OX NCBI_TaxID=2762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LB555 / PRINGSHEIM;
RX MEDLINE=91268060; PubMed=1711042;
RA Michalowski C.B., Loeffelhardt W., Bohnert H.J.;
RT "An ORF323 with homology to crtE, specifying prephytoene
RT pyrophosphate dehydrogenase, is encoded by cyanelle DNA in the
RT eukaryotic alga Cyanophora paradoxa.";
RT J. Biol. Chem. 266:11866-11870(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LB555 / PRINGSHEIM;
RA Stirewalt V.L., Michalowski C.B., Loeffelhardt W., Bohnert H.J.,
RA Bryant D.A.;
RT "Nucleotide sequence of the cyanelle DNA from Cyanophora paradoxa.";
RL plant Mol. Biol. Rep. 13:327-332(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=LB555 / PRINGSHEIM;
RA Loeffelhardt W., Stirewalt V.L., Michalowski C.B., Annarella M.,
RA Farley J.Y., Schlachter W.M., Chung S., Newmann-Spallart C.,
RA Steiner J.M., Jakowitsch J., Bohnert H.J., Bryant D.A.;
RT "The complete sequence of the cyanelle genome of Cyanophora paradoxa:
RT the genetic complexity of a primitive plastid.";
RL (In) Schenk H.E.A., Herrmann R., Jeon K.W., Mueller N.E.,
RL Schwemmler W. (eds.);
RL Eukaryotism and Symbiosis, pp.40-48, Springer-Verlag, Heidelberg
RL (1997).
RN [4]
RP SEQUENCE OF 1-46 FROM N.A.
RC STRAIN=LB555 / PRINGSHEIM;
RX MEDLINE=91117189; PubMed=2126059;
RA Michalowski C.B., Pfanzagl B., Loeffelhardt W., Bohnert H.J.;
RT "The cyanelle S10 spc ribosomal protein gene operon from Cyanophora
RT paradoxa.";
RL Mol. Gen. Genet. 224:222-231(1990).
CC -!- FUNCTION: POSSIBLE ROLE IN SYNTHESIS OF THE NONAPRENYL SIDE CHAIN
CC OF PLASTOQUINONE OR IN SYNTHESIS OF OTHER PRENYL CHAINS SUCH AS
CC UNDEKAPRENYL PYROPHOSPHATE.
CC -!- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.
CC -!- CAUTION: WAS ORIGINALLY (REF.1) CALLED CRTE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M37111; AAA65472.1; -;
DR EMBL; U30821; AAA81217.1; -;
DR EMBL; M30487; AAA63631.1; -;
DR PIR; A40433; A40433.
DR InterPro; IPR000092; Polyprenyl_synth.
DR Pfam; PF00348; polyprenyl_synth; 1.
DR PROSITE; PS00444; POLYPRENYL_SYNTHET_2; 1.
DR PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
KW Photosynthesis; Isoprene biosynthesis; Transferase; Cyanelle.
SQ SEQUENCE 323 AA; 35919 MW; CB4802466342B09A CRC64;

Query Match 31.6%; Score 43; DB 1; Length 323;
Best Local Similarity 42.1%; Pred. No. 19;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 5 ITXQCKENLPSCSXOGLR 23
DB 294 VAIQCLELPPSSSKALK 312

RESULT 18
YBP5_YEAST
ID YBP5_YEAST STANDARD; PRT; 418 AA.
AC P34220;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative deoxyribonuclease YBL055C (EC 3.1.21.-).
GN YBL055C OR YBL0512 OR YBL0511.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=94205266; PubMed=8154187;
RA Scherens B., el Bakkoury M., Vierendeels F., Dubois E., Messenguy F.;
RT "Sequencing and functional analysis of a 32,560 bp segment on the
RT left arm of yeast chromosome II. Identification of 26 open reading
RT frames, including the KIP1 and SEC17 genes.";
RL Yeast 9:1355-1371(1993).
CC -!- SIMILARITY: BELONGS TO THE TATD DNASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; J23261; CAAB0792.1; -;
DR EMBL; J35816; CAAB4875.1; -;
DR PIR; S39833; S39833.
DR PIR; S37334; S37334.
DR SGD; S0000151; YBL055C.
DR InterPro; IPR001130; TatD.
DR Pfam; PF01026; TatD_DNase; 1.
DR PROSITE; PS01137; TATD_1; FALSE_NEG.
DR PROSITE; PS01090; TATD_2; 1.
DR PROSITE; PS01091; TATD_3; 1.
KW Hypothetical protein; Hydrolase; Nuclease.
SQ SEQUENCE 418 AA; 47390 MW; F1107C1B8E94CAD1 CRC64;

Query Match 31.6%; Score 43; DB 1; Length 418;
Best Local Similarity 45.7%; Pred. No. 25;
Matches 16; Conservative 3; Mismatches 6; Indels 10; Gaps 3;

QY 1 TGSATYQCKENL-PSL-----CSXQGLRAENI 28
DB 280 TGSATDLOKLLNLSNIFIGVNGCS---LRTENL 311

RESULT 19
BIRG_MOUSE
ID BIRG_MOUSE STANDARD; PRT; 1402 AA.
AC Q9JTB3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Baculoviral IAP repeat-containing protein 19 (Neuronal apoptosis
DE inhibitory protein 7).
GN BIRC1G OR NAIP7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20414747; PubMed=10958627;
RA Endrizzi M.G., Hadinoto V., Gowney J.D., Miller W., Dietrich W.F.;
RT "Genomic sequence analysis of the mouse naip gene array.";
RL Genome Res. 10:1095-1102(2000).
CC -!- FUNCTION: PREVENTS MOTOR-NEURON APOPTOSIS INDUCED BY A VARIETY OF
CC SIGNALS.
CC -!- SIMILARITY: CONTAINS 3 BIR REPEATS.
CC -!- SIMILARITY: CONTAINS 1 NACHT DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AF242433; AAF82749.1; -;
DR MGD; MGI:1858256; Birc1g.
DR InterPro; IPR001370; BIR.
DR Pfam; PF00653; BIR; 3.
DR SMART; SM00238; BIR; 3.
DR PROSITE; PS01282; BIR_REPEAT_1; 2.
DR PROSITE; PS01443; BIR_REPEAT_2; 3.
DR PROSITE; PS50837; NACHT; 1.
KW Apoptosis; Repeat; Multigene family.
FT REPEAT 60 127 BIR 1.
FT REPEAT 159 227 BIR 2.
FT REPEAT 278 345 BIR 3.
FT DOMAIN 464 759 NACHT.
SQ SEQUENCE 1402 AA; 159662 MW; C1DFFBA359893E0D CRC64;

Query Match 31.6%; Score 43; DB 1; Length 1402;
Best Local Similarity 50.0%; Pred. No. 84;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 6 TXQCKENLPSCSXOG 21
DB 414 TSFCHNMLPEVCSSLG 429

RESULT 20
BIRF_MOUSE
ID BIRF_MOUSE STANDARD; PRT; 1403 AA.
AC Q9JTB6; P81704; O09122; O09121;
DT 16-OCT-2001 (Rel. 40, Created)


```

GN STC2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE=96389283; PubMed=9723890;
RX Chang A.C.-M., Reddel R.R.;
RT "Identification of a second stanniocalcin cDNA in mouse and human:
RT stanniocalcin 2.";
RL Mol. Cell. Endocrinol. 141:95-99(1998).
RN [2]
RN SEQUENCE FROM N.A.
RP TISSUE=Osteosarcoma;
RX Ishibashi K., Miyamoto K., Taketani Y., Morita K., Takeda E.,
RA Sasaki S., Imai M.;
RT "Molecular cloning of a second human stanniocalcin homologue (STC2).";
RL Biochem. Biophys. Res. Commun. 250:252-258(1998).
RN [3]
RN SEQUENCE FROM N.A.
RX DIMattia G.E., Varghese R., Wagner G.F.;
RT "Molecular cloning and characterization of stanniocalcin-related
RT protein.";
RL Mol. Cell. Endocrinol. 146:137-140(1998).
RN [4]
RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX TISSUE=Breast carcinoma;
RP MEDLINE=99145369; PubMed=10022771;
RX DIMattia G.E., Varghese R., Wagner G.F.;
RT "Molecular cloning and characterization of stanniocalcin-related
RT protein.";
RL Mol. Cell. Endocrinol. 146:137-140(1998).
CC [-] FUNCTION: HAS AN ANTI-HYPOCALCEMIC ACTION ON CALCIUM AND PHOSPHATE
CC HOMEOSTASIS.
CC [-] SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC [-] SUBCELLULAR LOCATION: Secreted (Potential).
CC [-] TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES INCLUDING
CC MUSCLE, HEART, PANCREAS, KIDNEY, SPLEEN, PROSTATE, SMALL
CC INTESTINE, COLON AND PERIPHERAL BLOOD LEUCOCYTES.
CC [-] SIMILARITY: BELONGS TO THE STANNIOCALCIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF055460; AAC27036.1; -
DR EMBL; AB012664; BAA33489.1; -
DR EMBL; AF098462; AAC97948.1; -
DR EMBL; AF031036; AAD01922.1; -
DR MIM; 603665; -
KW Hormone; Signal; Glycoprotein.
FT SIGNAL 1 24
FT CHAIN 25 302
FT CARBOHYD 73 73
FT SEQUENCE 302 AA; 33248 MW; 9B90D8911524FA22 CRC64;
FT STANNIOCALCIN 2.
FT N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 302 AA; 33248 MW; 9B90D8911524FA22 CRC64;
-----
Query Match 30.9%; Score 42; DB 1; Length 302;
Best Local Similarity 46.7%; Pred. No. 26;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 4 AITXCKENLPGLS 18
Db 192 SVQOCEQNWGLS 206
:: |||:| ||||

```

```

RESULT 24
ID STC2_MACNE STANDARD; PRT; 302 AA.
AC O97561; 2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Stanniocalcin 2 precursor (STC-2).
GN STC2
OS Macaca nemestrina (Pig-tailed macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9545;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreatic islets;
RX MEDLINE=99378040; PubMed=10450831;
RA Moore E.E., Kuestner R.E., Conklin D.C., Whitmore T.E., Downey W.,
RA Buddle M.M., Adams R.L., Bell L.A., Thompson D.L., Wolf A., Chen L.,
RA Stamm M.R., Grant F.J., Lok S., Ren H., de Jongh K.S.;
RT "Stanniocalcin 2: characterization of the protein and its localization
RT to human pancreatic alpha cells.";
RL Horm. Metab. Res. 31:406-414(1999).
CC -!- FUNCTION: HAS AN ANTI-HYPOCALCEMIC ACTION ON CALCIUM AND PHOSPHATE
CC HOMIOSTASIS.
CC -!- SUBUNIT: HOMODIMER: DISULFIDE-LINKED (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- SIMILARITY: BELONGS TO THE STANNIOCALCIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF035377; AAD02027.1; -.
CC Hormone; Signal; Glycoprotein.
CC SIGNAL 1 24 POTENTIAL.
CC CHAIN 25 302 STANNIOCALCIN 2.
CC CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 302 AA; 33281 MW; 62953CB958AF64C1 CRC64;
CC -----
Query Match 30.9%; Score 42; DB 1; Length 302;
Best Local Similarity 46.7%; Pred. No. 26;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 4 AITXQCKENLPSCLS 18
DB 192 SVQVQCEQNWGSLCS 206
:: ||::| ||||
-----
RESULT 25
ID VMAT_SV5 STANDARD; PRT; 377 AA.
AC P16629;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Matrix protein (Membrane protein).
GN M.
OS Simian virus 5 (strain W3) (SV5).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Rubulavirus.
OX NCBI_TaxID=11208;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90322733; PubMed=2330672;
-----
Query Match 30.9%; Score 42; DB 1; Length 302;
Best Local Similarity 46.7%; Pred. No. 26;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 4 AITXQCKENLPSCLS 18
DB 192 SVQVQCEQNWGSLCS 206
:: ||::| ||||
-----
RA Sheshberadaran H., Lamb R.A.;
RT "Sequence characterization of the membrane protein gene of
RT paramyxovirus simian virus 5.";
RL Virology 176:234-243(1990).
CC -!- FUNCTION: THE M PROTEIN HAS A CRUCIAL ROLE IN VIRUS ASSEMBLY
CC AND INTERACTS WITH THE RNP COMPLEX AS WELL AS WITH THE VIRAL
CC MEMBRANE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M32248; AAA46901.1; -.
CC EMBL; AF052755; AAC95514.1; -.
CC PIR; A34604; MENZS5.
CC InterPro; IPR000982; Matrix.
CC Pfam; PF00661; Matrix; 1.
CC ProDom; PD000741; Matrix; 1.
CC Matrix protein; Envelope protein.
CC SEQUENCE 377 AA; 42250 MW; DF424183C87887C3 CRC64;
CC -----
Query Match 30.9%; Score 42; DB 1; Length 377;
Best Local Similarity 42.9%; Pred. No. 32;
Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
QY 1 TGSAITXQCKENLPSCLSXOG 21
DB 85 TAAMVTLCGCPNLPGLNVIG 105
:::|:| ||||:|
-----
RESULT 26
ID VG39_HSV11 STANDARD; PRT; 1123 AA.
AC Q00143;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Hypothetical gene 39 protein.
GN 39.
OS Ictalurid herpesvirus 1 (Channel catfish virus) (CCV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC unclassified Herpesviridae.
OX NCBI_TaxID=10401;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AUBURN 1;
RX MEDLINE=92087490; PubMed=1727613;
RA Davison A.J.;
RT "Channel catfish virus: a new type of herpesvirus.";
RL Virology 186:9-14(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M75136; AAA88142.1; -.
CC PIR; D36790; D36790.
CC Hypothetical protein.
CC SEQUENCE 1123 AA; 123008 MW; 47953EBE727EBD00 CRC64;
CC -----
Query Match 30.5%; Score 41.5; DB 1; Length 1123;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

```



```
CC EMBL; S79538; -; NOT_ANNOTATED_CDS.
DR EMBL; S79537; -; NOT_ANNOTATED_CDS.
DR EMBL; AA220754; -; NOT_ANNOTATED_CDS.
DR EMBL; AA220626; -; NOT_ANNOTATED_CDS.
DR EMBL; AA015403; -; NOT_ANNOTATED_CDS.
DR MGD; MGI:1202875; Btf3.
DR pfam; PF01849; NAC; 1.
KW Transcription regulation; Nuclear protein; Alternative splicing.
FT CHAIN 1 204 BTF3A (BY SIMILARITY).
FT CHAIN 43 204 BTF3B (BY SIMILARITY).
FT CHAIN 47 47 I -> V (IN REF. 1).
FT CONFLICT 55 55 K -> E (IN REF. 2; AA015403).
FT CONFLICT 55 55 K -> R (IN REF. 1).
FT CONFLICT 109 109 N -> S (IN REF. 1).
FT CONFLICT 149 150 LN -> RH (IN REF. 2; AA015403).
SQ SEQUENCE 204 AA; 21999 MW; 78318EC1E01C9523 CRC64;

Query Match 30.1%; Score 41; DB 1; Length 204;
Best Local Similarity 47.6%; Pred. No. 25;
Matches 10; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 TGSATYXQCKENLPSCSXQG 21
||| | | | | | | | | | | | | | |
DB 133 TGAETKQLTEMLPSILNQLG 153

RESULT 29
BTF3_HUMAN
ID BTF3_HUMAN STANDARD; PRT; 206 AA.
AC P20290; Q13893;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription factor BTF3 (RNA polymerase B transcription factor 3).
GN BTF3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90206080; PubMed=2320128;
RA Zheng X.M., Black D., Chambon P., Egly J.-M.;
RT "Sequencing and expression of complementary DNA for the general
transcription factor BTF3.";
RL Nature 344:556-559(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Leukocyte;
RX MEDLINE=92347696; PubMed=1386332;
RA Kanno M., Chalut C., Egly J.-M.;
RT "Genomic structure of the putative BTF3 transcription factor.";
RL Gene 117:219-228(1992).
CC -!- FUNCTION: GENERAL TRANSCRIPTION FACTOR. BTF3 CAN FORM A STABLE
COMPLEX WITH RNA POLYMERASE II. REQUIRED FOR THE INITIATION OF
TRANSCRIPTION.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; BTF3A (SHOWN HERE) AND BTF3B;
ARE PRODUCED BY ALTERNATIVE SPLICING. BTF3B IS TRANSCRIPTIONALLY
INACTIVE DESPITE ITS ABILITY TO BIND RNAP II.
CC -!- SIMILARITY: BELONGS TO THE BTF3 TRANSCRIPTION FACTORS FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```

```
DR EMBL; X53280; CAA37375.1; -.
DR EMBL; M90357; AAA58398.1; -.
DR EMBL; M90352; AAA58398.1; JOINED.
DR PIR; S09160; S09160.
DR PIR; S09161; S09161.
DR MIM; 602542; -.
DR InterPro; IPR002715; NAC.
DR pfam; PF01849; NAC; 1.
KW Transcription regulation; Nuclear protein; Alternative splicing.
FT CHAIN 1 206 BTF3A.
FT CHAIN 45 206 BTF3B.
FT CHAIN 41 41 Q -> E (IN REF. 2).
FT CONFLICT 68 105 MISSING (IN REF. 2).
FT CONFLICT 192 196 DLVEN -> GG (IN REF. 2).
FT CONFLICT 198 198 D -> Q (IN REF. 2).
SQ SEQUENCE 206 AA; 22168 MW; 9653AC480EAF64C6 CRC64;

Query Match 30.1%; Score 41; DB 1; Length 206;
Best Local Similarity 47.6%; Pred. No. 25;
Matches 10; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 TGSATYXQCKENLPSCSXQG 21
||| | | | | | | | | | | | | | |
DB 135 TGAETKQLTEMLPSILNQLG 155

RESULT 30
YTEA_BACSU
ID YTEA_BACSU STANDARD; PRT; 239 AA.
AC P42408;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 27.7 kDa protein (ORFQ).
GN YTEA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RBI;
RA Qin X., Taber H.W.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: IN THE N-TERMINAL HALF, TO THE N-TERMINAL OF
E. COLI DKSA.
CC -!- SIMILARITY: SOME, TO B. SUBTILIS YLYA AND YOCK.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```

RESULT 31

```
VP10_WTV
ID VP10_WTV STANDARD; PRT; 345 AA.
AC P13093;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-AUG-1991 (Rel. 19, Last annotation update)
DE Nonstructural protein PMS10.
GN S9.
OS Wound tumor virus (WTV).
OC Viruses; dsRNA viruses; Reoviridae; Phytoreovirus.
OX NCBI_TaxID=10987;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89299461; PubMed=2525838;
RA Anzola J.V., Dall D.J., Xu Z., Nuss D.L.;
RT "Complete nucleotide sequence of wound tumor virus genomic segments
RL encoding nonstructural polypeptides.";
RL Virology 171:222-228(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M24115; AAA48504.1; -
DR PIR; C32442; MNXRWT.
KW Nonstructural protein.
SQ SEQUENCE 345 AA; 38610 MW; FE866FAE39A39201 CRC64;

Query Match 30.1%; Score 41; DB 1; Length 345;
Best Local Similarity 40.0%; Pred No. 42;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 8 QCKENLPSLC SXOGL 22
DB 112 QCNSTPSPCNSEGI 126

RESULT 32
11SB_CUCMA STANDARD; PRT; 480 AA.
AC P13744;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DE 11S globulin beta subunit precursor.
OS Cucurbita maxima (Pumpkin) (Winter squash).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.
OX NCBI_TaxID=3661;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. KUOKAWA AMAKURI NANKIN;
RX MEDLINE=88166744; PubMed=2450746;
RA Hayashi M., Mori H., Nishimura M., Akazawa T., Hara-Nishimura I.;
RT "Nucleotide sequence of cloned cDNA coding for pumpkin 11-S globulin
RL beta subunit.";
RL Eur. J. Biochem. 172:627-632(1988).
RN [2]
RP SEQUENCE OF 22-30 AND 297-302.
RA Ohmura M., Hara I., Mastubara H.;
RT "pumpkin (Cucurbita sp.) seed globulin IV. Terminal sequences of the
RT acidic and basic peptide chains and identification of a pyroglutamyl
RT peptide chain.";
RL Plant Cell Physiol. 21:157-167(1980).
CC -1- FUNCTION: THIS IS A SEED STORAGE PROTEIN.
CC -1- SUBUNIT: HEXAMER; EACH SUBUNIT IS COMPOSED OF AN ACIDIC AND A
CC BASIC CHAIN DERIVED FROM A SINGLE PRECURSOR AND LINKED BY A
```

```
CC DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE 11S SEED STORAGE PROTEIN (GLOBULINS)
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M36407; AAA33110.1; -
DR PIR; S00366; FWPUIB.
DR InterPro; IPR000459; Seedstore_11s.
DR Pfam; PF00190; Seedstore_11s; 1.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S_SEED_STORAGE; 1.
KW Seed storage protein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 480 11S GLOBULIN BETA SUBUNIT.
FT CHAIN 22 296 GAMMA CHAIN (ACIDIC).
FT CHAIN 257 480 DELTA CHAIN (BASIC).
FT MOD_RES 22 22 PYRROLIDONE CARBOXYLIC ACID.
FT DISULFID 124 303 INTERCHAIN (GAMMA-DELTA) (POTENTIAL).
FT CONFLICT 27 27 S -> E (IN REF. 2).
FT CONFLICT 30 30 E -> S (IN REF. 2).
SQ SEQUENCE 480 AA; 54625 MW; BCD8A83DD1AED93C CRC64;

Query Match 30.1%; Score 41; DB 1; Length 480;
Best Local Similarity 29.6%; Pred. No. 59;
Matches 8; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 2 GSATYQCKENLPSLC SXOGLRAEENI 28
DB 33 GSEVMQOHHRYQSPRACRLENLRAQDPV 59

RESULT 33
CCB2_DROME STANDARD; PRT; 561 AA.
AC Q9V8M2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable cytochrome P450 12b2, mitochondrial precursor (EC 1.14.-.-)
DE (CYP11B2).
GN CYP12B2 OR CG15077.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Fabros B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
```

```

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kimms I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
RT Science 287:2185-2195(2000).
RL
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Mitochondrial (Potential).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE003798; AAF57642.1; -.
DR FlyBase; FBgn0034387; Cyp12b2.
DR InterPro; IPR001128; Cyt_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
DR Oxidoreductase; Monooxygenase; Membrane; Heme; Mitochondrion;
KW Transit peptide; Hypothetical protein.
FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
FT CHAIN ? 561 PROBABLE CYTOCHROME P450 12B2.
FT BINDING 505 505 HEME (BY SIMILARITY)
FT BINDING 561 AA; 64601 MW; 4135ABB3D331E565 CRC64;
SQ SEQUENCE 561 AA; 64601 MW; 4135ABB3D331E565 CRC64;

Query Match 30.1%; Score 41; DB 1; Length 561;
Best Local Similarity 40.0%; Pred. No. 69;
Matches 10; Conservative 5; Mismatches 8; Indels 2; Gaps 1;

QY 1 TGSATYXCKENLPSL--CSXQGLR 23
DQ 394 TRKSLTDQNTKNMPLRACIKEGLR 418

RESULT 34
SYFB_SYNP7
ID SYFB_SYNP7 STANDARD; PRT; 808 AA.
AC P4764;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20) (Phenylalanine--
DE tRNA ligase beta chain) (PheRS).
OS PHE.
GN Synchococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.
OX NCBI_TaxID=1140;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA Cranenburgh R.M., Robinson N.J.;
RT "Phenylalanyl-tRNA synthetase gene, pheT, from Synchococcus PCC
RT 7942.";
RL J. Appl. Phycol. 8:81-82(1996).
CC -!- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP +
CC diphosphate + L-phenylalanyl-tRNA(Phe).
CC -!- SUBUNIT: Tetramer of two alpha and two beta chains (by
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE PHENYLALANYL-TRNA SYNTHETASE BETA
CC CHAIN FAMILY. SUBFAMILY 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X94345; CAA64071.1; -.
DR HSP; P27002; LPYS.
DR InterPro; IPR002547; tRNA_bind.
DR Pfam; PF01588; tRNA_bind; 2.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
SQ SEQUENCE 808 AA; 89840 MW; 501BBEAC44624A6A CRC64;

Query Match 30.1%; Score 41; DB 1; Length 808;
Best Local Similarity 66.7%; Pred. No. 99;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 14 PSLCSXQGLRAE 25
DQ 674 PSLCEGRGLPAE 685

RESULT 35
YE20_METJA
ID YE20_METJA STANDARD; PRT; 1102 AA.
AC Q58815;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ1420 [Contains: Mja gf6p intein].
GN MJ1420.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
CC -!- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
CC A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)
CC FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).
CC -!- SIMILARITY: IN THE C-TERMINAL, TO GLUCOSAMINE--FRUCTOSE-6-
CC PHOSPHATE AMINOTRANSFERASES (GFAT).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

KA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,

DR HSSP; P07272; IPYI.
DR SGD; S0000444; THI2.

RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman

DR HSSP; P07272; IPYI.
DR SGD; S0000444; THI2.

DR Pfam; PF00172; Zn_clus; 1.
DR SMART; SM00066; GAL4; 1.
DR PROSITE; PS00463; ZN2_CY6_FUNGAL_1; 1.
DR PROSITE; PS50048; ZN2_CY6_FUNGAL_2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein; Zinc;
KW Metal-binding; Activator; Thiamine biosynthesis.
FT DNA_BIND 30 57 ZN(2)-CYS(6), FUNGAL-TYPE.
SQ SEQUENCE 450 AA; 51822 MW; 1A8B1627EC66FD93 CRC64;

Query Match 29.8%; Score 40.5; DB 1; Length 450;
Best Local Similarity 37.5%; Pred. No. 66;
Matches 12; Conservative 3; Mismatches 6; Indels 11; Gaps 2;

QY 8 QCKENLP--SICSGQG-----LRAEENI 28
:| | | | | | | | | | | | | | | | | | | |
DB 39 RCDENRPFICSLCAKHGDCSYDIRLWLEENI 70
:| | | | | | | | | | | | | | | | | | | |

RESULT 38
ANGI_MOUSE STANDARD; PRT; 145 AA.
AC P21570;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Angiogenin precursor (EC 3.1.27.-).
GN ANG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91025023; PubMed-2222458;
RA Bond M.D., Vallee B.L.;
RT "Isolation and sequencing of mouse angiogenin DNA.";
RL Biochem. Biophys. Res. Commun. 171:988-995(1990).
RN [2]
RP PARTIAL SEQUENCE.
RC TISSUE=Serum;
RX MEDLINE-93192291; PubMed-8448182;
RA Bond M.D., Strydom D.J., Vallee B.L.;
RT "Characterization and sequencing of rabbit, pig and mouse
RT angiogenins: discernment of functionally important residues and
RT regions.";
RL Biochim. Biophys. Acta 1162:177-186(1993).
CC -!- FUNCTION: MAY FUNCTION AS A TRNA-SPECIFIC RIBONUCLEASE THAT BINDS
CC TO ACTIN ON THE SURFACE OF ENDOTHELIAL CELLS; ONCE BOUND,
CC ANGIOGENIN IS ENDOCYTOSED AND TRANSLOCATED TO THE NUCLEUS, THEREBY
CC PROMOTING THE ENDOTHELIAL INVASIVENESS NECESSARY FOR BLOOD VESSEL
CC FORMATION. ANGIOGENIN INDUCES VASCULARIZATION OF NORMAL AND
CC MALIGNANT TISSUES. ABOLISHES PROTEIN SYNTHESIS BY SPECIFICALLY
CC HYDROLYZING CELLULAR TRNAS.
CC -!- SIMILARITY: BELONGS TO THE PANCREATIC RIBONUCLEASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U22516; AAA91366.1; -.
CC PIR; A35932; A35932.
CC HSSP; P03950; 1A4Y.
CC MGD; MGI:88022; Ang.
CC InterPro; IPR001427; RnaseA.
CC Pfam; PF00074; rnaseA; 1.
CC PRINTS; PR00794; RIBONUCLEASE.
CC ProDom; PD000535; RnaseA; 1.
CC SMART; SM00092; RNase_Pc; 1.

DR PROSITE; PS00127; RNASE_PANCREATIC; 1.
KW Hydrolase; Nuclease; Endonuclease; Angiogenesis;
KW Protein synthesis inhibitor; Signal.
FT SIGNAL 1 24
FT CHAIN 25 145 ANGIOGENIN.
FT MOD_RES 25 25 PYRROLIDONE CARBOXYLIC ACID
FT ACT_SITE 37 37 (BY SIMILARITY).
FT ACT_SITE 64 64 BY SIMILARITY.
FT ACT_SITE 137 137 BY SIMILARITY.
FT DISULFID 50 104 BY SIMILARITY.
FT DISULFID 63 115 BY SIMILARITY.
FT DISULFID 81 130 BY SIMILARITY.
SQ SEQUENCE 145 AA; 16228 MW; 06944260BB764938 CRC64;

Query Match 29.4%; Score 40; DB 1; Length 145;
Best Local Similarity 22.9%; Pred. No. 26;
Matches 8; Conservative 7; Mismatches 10; Indels 10; Gaps 1;

QY 4 ALTXQCKE-----NLPLSLCSXQGLRAEENI 28
:| | | | | | | | | | | | | | | | | | | |
DB 58 SLTSPCKDVNTFIHGKNSNIKAICGANGSPYRENL 92
:| | | | | | | | | | | | | | | | | | | |

RESULT 39
NEU2_CATCO STANDARD; PRT; 148 AA.
ID NEU2_CATCO
AC P15211;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Isotocin-neurophysin IT 2 precursor [Contains: Isotocin (IT);
DE Neurophysin IT 2].
OS Catostomus commersoni (White sucker).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Catostomidae; Catostomus.
OX NCBI_TaxID=7971;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-90059870; PubMed-2583084;
RA Figueroa J., Morley S.D., Heierhorst J., Krentler C., Lederis K.,
RA Richter D.;
RT "Two isotocin genes are present in the white sucker Catostomus
RT commersoni both lacking introns in their protein coding regions.";
RL EMBO J. 8:2873-2877(1989).
CC -!- FUNCTION: ISOTOCIN CAUSES CONTRACTION OF SMOOTH MUSCLES.
CC -!- PTM: SEVEN DISULFIDE BONDS ARE PRESENT IN NEUROPHYSIN.
CC -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X16622; CAA34619.1; -.
CC PIR; S06019; S06019.
CC HSSP; P01180; INPO.
CC InterPro; IPR000981; Neurohypophys_horm.
CC Pfam; PF00220; hormone4; 1.
CC Pfam; PF00184; hormone5; 1.
CC PRINTS; PR00831; NEUROPHYSIN.
CC ProDom; PD001676; Neurohypophys_horm; 1.
CC SMART; SM00003; NH; 1.
CC KW Hormone; Hypothalamus; Amidation; Cleavage on pair of basic residues;
KW Hormone; Hypothalamus; Amidation; Cleavage on pair of basic residues;
KW Multigene family; Signal.
FT SIGNAL 1 20
FT PEPTIDE 21 29 ISOTOCIN.

```
FT PEPTIDE      32 148      NEUROPHYSIN IT 2.
FT DISULFID    21 26
FT MOD_RES     29 29      AMIDATION (G-30 PROVIDE AMIDE GROUP).
SQ SEQUENCE    148 AA; 15581 MW; 26A9EDC2513E9A6F CRC64;

Query Match      29.4%; Score 40; DB 1; Length 148;
Best Local Similarity 45.5%; Pred. No. 26;
Matches 10; Conservative 2; Mismatches 8; Indels 2; Gaps 1;

QY 2 GSAITXOCKEN--LPSLCSXQG 21
   || | :|:| ||| | |
Db 69 GSPETLRQOEDEFLPSPCEAGG 90

RESULT 40
STC2_CAVPO STANDARD; PRT; 197 AA.
AC P57675;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Stanniocalcin 2 (STC-2) (Fragments).
GN STC2.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE.
RX MEDLINE=99378040; PubMed=10450831;
RA Moore E.E., Kuestner R.E., Conklin D.C., Whitmore T.E., Downey W.,
RA Buddle M.M., Adams R.L., Bell L.A., Thompson D.L., Wolf A., Chen L.,
RA Stamm M.R., Grant F.J., Lok S., Ren H., de Jongh K.S.;
RT "Stanniocalcin 2: characterization of the protein and its localization
RT to human pancreatic alpha cells.";
RL Horm. Metab. Res. 31:406-414(1999).
CC -|- FUNCTION: HAS AN ANTI-HYPOCALCEMIC ACTION ON CALCIUM AND PHOSPHATE
CC HOMEOSTASIS.
CC -|- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC -|- SUBCELLULAR LOCATION: Secreted (Potential).
CC -|- SIMILARITY: BELONGS TO THE STANNIOCALCIN FAMILY.
KW Hormone.
FT NON_CONS    60 61
FT NON_CONS    64 65
FT NON_CONS    69 70
FT NON_CONS    75 76
FT NON_CONS    92 93
FT NON_CONS    154 155
SQ SEQUENCE    197 AA; 21899 MW; 1512DA859C8A8E67 CRC64;
```

```
Query Match      29.4%; Score 40; DB 1; Length 197;
Best Local Similarity 46.7%; Pred. No. 35;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 AITXOCKENLPSLCS 18
   :| ||:| |||
Db 126 SIQAOCEONWGLCS 140
```

Search completed: August 22, 2002, 07:52:33
Job time: 191 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 07:52:14 ; Search time 41.85 Seconds
(without alignments)
115.743 Million cell updates/sec

Title: US-09-856-199-2

Perfect score: 136

Sequence: 1 TGAATYQCKENLPSCSXOGLRAENI 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

SPTREMBL_19.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rhodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriaph.*

17: sp_archaeop.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	80.9	386	6 Q9N2H6	Q9n2h6 bos taurus
2	102	75.0	380	11 Q9Z2T7	Q9z2t7 mus musculus
3	52	38.2	670	5 Q9W7V6	Q9w7v6 drosophila
4	52	38.2	679	5 Q9V7V7	Q9v7v7 drosophila
5	51.5	37.9	219	4 Q9H4C9	Q9h4c9 homo sapien
6	51.5	37.9	1624	4 Q9H4C9	Q9h4c9 homo sapien
7	51.5	37.9	1790	4 Q9BXP8	Q9bxp8 homo sapien
8	51.5	37.9	1791	4 Q9BXP8	Q9bxp8 homo sapien
9	51	37.5	151	16 Q9HTJ9	Q9htj9 pseudomonas
10	50.5	37.1	2931	5 Q9W2C6	Q9w2c6 drosophila
11	47	34.6	358	5 O45893	O45893 caenorhabdi
12	47	34.6	1207	4 O75136	O75136 homo sapien
13	46	33.8	146	5 O17233	O17233 caenorhabdi
14	46	33.8	601	11 Q9DBR2	Q9dbrr2 mus musculus
15	46	33.8	1002	10 Q80875	Q80875 arabidopsis
16	46	33.8	1002	10 Q94BS5	Q94bs5 arabidopsis

Q9dat2 mus musculus
Q9cn47 pasteurilla
Q9fgn6 arabidopsis
Q9kg25 bacillus ha
Q9d3b8 mus musculus
Q9dnt6 porcine rub
Q13792 schizosacch
Q9pez3 xylella fas
Q9bq17 homo sapien
Q9y5d0 homo sapien
Q9y5g6 homo sapien
Q962j8 plasmodium
Q97354 spodoptera
Q95935 pyrococcus
P97337 mus musculus
Q9c483 pneumocysti
Q9w5p0 drosophila
Q9mjc3 zea mays (m
Q98sv1 brachydanio
Q9lfn6 arabidopsis
Q93vj8 arabidopsis
Q9lfn9 arabidopsis
P74623 synecocyst
Q98603 peste-des-p
Q93cg7 photobacter
Q9aus0 oryza sativ
Q94h07 oryza sativ
Q9x5j5 mycobacteri
Q9pk0 homo sapien
Q9xw5 caenorhabdi
Q9avk2 lycopersico
Q9xyy6 caenorhabdi
O56941 human papil
P89953 hepatitis c
Q9vnf4 drosophila
Q93dt2 streptococc
Q9k478 streptomyce
Q9hxx1 pseudomonas
Q9gw84 rat cytomeg
Q57079 cowpox viru
Q72761 cowpox viru
Q9exx0 mus musculus
Q39664 hepatitis c
Q9gu80 trypanosoma
Q9pgp5 trypanosoma
Q9hb96 homo sapien
Q21740 caenorhabdi
Q9c9r6 arabidopsis
Q9fnc7 arabidopsis
Q941l4 arabidopsis
Q9vcs3 drosophila
Q962i0 drosophila
Q9xu67 caenorhabdi
Q9ub90 homo sapien
Q9y3p9 mus musculus
Q64449 mus musculus
Q9nwe9 homo sapien
Q9iu72 human immun
Q9klh5 neisseria m
Q9jx24 neisseria m
Q9h7v5 homo sapien
Q96st4 homo sapien
Q9p217 homo sapien
Q9zew7 klebsiella
Q96o69 homo sapien
Q91101 human immun
Q9yhp3 ginglymosto
Q65610 arabidopsis
Q43589 nicotiana t
Q9u9b2 ancylostoma
Q23108 arabidopsis
Q9sve5 arabidopsis
Q9fza4 arabidopsis

90 41 30.1 316 5 Q96729 giardia lam
91 41 30.1 348 17 Q96YJ3 sulfolobus
92 41 30.1 445 16 Q91012 pseudomonas
93 41 30.1 452 10 Q91X41 arabidopsis
94 41 30.1 459 10 Q949V7 arabidopsis
95 41 30.1 481 15 Q91V19 arabidopsis
96 41 30.1 527 10 Q9STD7 arabidopsis
97 41 30.1 571 2 Q9LBE1 arabidopsis
98 41 30.1 571 2 Q9KH5 arabidopsis
99 41 30.1 607 10 Q82223 arabidopsis
100 41 30.1 641 2 Q9XCC8 streptomyces

ALIGNMENTS

RESULT 1
Q9N2H6 PRELIMINARY; PRT; 386 AA.
AC Q9N2H6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE SELENOPROTEIN P.
GN SELP.

OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Hara S., Imura N., Shoji Y.
RT "Bovine endothelial selenoprotein P."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB032826; BAA84781.1; --
SQ SEQUENCE 386 AA; 43061 MW; EC6F3A76874E528F CRC64;

Query Match 80.9%; Score 110; DB 6; Length 386;
Best Local Similarity 78.6%; Pred. No. 2.4e-10;
Matches 22; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TGSATYQCKENLPSCSXQGLRAENI 28
DB 316 TGSATTCQCTEKLPSCSQGLAEENV 343

RESULT 2
Q922T7 PRELIMINARY; PRT; 380 AA.
AC Q922T7;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE PLASMA SELENOPROTEIN P.
GN SELP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Steinert P., Ahrens M., Gross G., Flohe L.;
RT "cDNA and deduced polypeptide sequence of a mouse selenoprotein P."
RL BioFactors 6:311-319(1997).
DR EMBL: AF021345; AAD01684.1; --
SQ SEQUENCE 380 AA; 42375 MW; C5055C18D32BDD14 CRC64;

Query Match 75.0%; Score 102; DB 11; Length 380;

Best Local Similarity 78.6%; Pred. No. 5.4e-09;
Matches 22; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 TGSATYQCKENLPSCSXQGLRAENI 28
DB 312 SGSAITAXQCAENLPSCSXQGLFAEEKV 339

RESULT 3
Q9V7V6 PRELIMINARY; PRT; 670 AA.
AC Q9V7V6;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE CG6622 PROTEIN.
GN PKC33E OR CG6622.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S.D., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Durbin C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AE003805; AAF57932.1; --
DR HSSP: P05697; ITBN.
DR FlyBase: FBgn0003091; Pkc53E.
DR InterPro: IPR000008; C2.
DR InterPro: IPR002219; DAG_PE-bind.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR000961; Pkinase_C.
DR InterPro: IPR002290; Ser_thr_pkinase.

DR pfam: PF00168; C2; 1.
 DR pfam: PF00130; DAG_PE-bind; 2.
 DR pfam: PF00069; pkinase; 1.
 DR pfam: PF00433; pkinase_C; 1.
 DR PRINTS; PR00360; C2DOMAIN.
 DR PRINTS; PR00008; DAGPEDOMAIN.
 DR SMART; SM00109; C1; 2.
 DR SMART; SM00239; C2; 1.
 DR SMART; SM00220; S_TK_X; 1.
 DR SMART; SM00133; S_TK_X; 1.
 DR PROSITE; PS50004; C2_DOMAIN.2; 2.
 DR PROSITE; PS00479; DAG_PE_BIND_DOM.1; 2.
 DR PROSITE; PS50081; DAG_PE_BIND_DOM.2; 4.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 2.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 670 AA; 76733 MW; 588F5194A212BBAF CRC64;

Query Match 38.2%; Score 52; DB 5; Length 670;
 Best Local Similarity 61.5%; Pred. No. 2.9;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 ITXOCKENLPSLC 17
 : : |||||:||||
 Db 148 VHARCKENVPSLC 160

RESULT 4
 ID Q9V7V7 PRELIMINARY; PRT; 679 AA.
 AC Q9V7V7;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE CG6622 PROTEIN
 GN PKC33E OR CG6622.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 ON NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hostkins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Fandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Brill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Chertys J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AE003805; AAF57933.1; -;
 DR HSSP; P05697; ITBN.
 DR FlyBase; FBgn0003091; Pkc53E.
 DR InterPro; IPR000008; C2.
 DR InterPro; IPR002219; DAG_PE-bind.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR000961; Pkinase_C.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR Pfam; PF00168; C2; 1.
 DR Pfam; PF00130; DAG_PE-bind; 2.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF00433; pkinase_C; 1.
 DR PRINTS; PR00360; C2DOMAIN.
 DR PRINTS; PR00008; DAGPEDOMAIN.
 DR SMART; SM00109; C1; 2.
 DR SMART; SM00239; C2; 1.
 DR SMART; SM00220; S_TK_X; 1.
 DR SMART; SM00133; S_TK_X; 1.
 DR PROSITE; PS00499; C2_DOMAIN.1; 1.
 DR PROSITE; PS50004; C2_DOMAIN.2; 1.
 DR PROSITE; PS00479; DAG_PE_BIND_DOM.1; 1.
 DR PROSITE; PS50081; DAG_PE_BIND_DOM.2; 2.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 679 AA; 77694 MW; 3C69AD351E36B7DC CRC64;

Query Match 38.2%; Score 52; DB 5; Length 679;
 Best Local Similarity 61.5%; Pred. No. 2.9;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 ITXOCKENLPSLC 17
 : : |||||:||||
 Db 157 VHARCKENVPSLC 169

RESULT 5
 ID Q75997 PRELIMINARY; PRT; 219 AA.
 AC Q75997;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE DJ774124.1 (SIMILAR TO PREGNANCY-ASSOCIATED PLASMA PROTEIN A
 DE PRECURSOR) (FRAGMENT).
 GN DJ774124.1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ON NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Grafham D.;
 RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL031290; CAA20358.1; -;
 DR InterPro; IPR000800; Notch.
 DR InterPro; IPR000436; Sushi_SCR_CCP.
 DR SMART; SM00032; CCP; 1.
 DR SMART; SM00004; NL; 1.


```

RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Coltray L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.A., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Salier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004947; AAC08744.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 151 AA; 16472 MW; B7A788408A09B49C CRC64;

Query Match 37.5%; Score 51; DB 16; Length 151;
Best Local Similarity 45.5%; Pred. No. 0.87;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 3 SAITXOCKENLPSCSXOGLRA 24
II : II : II : II : II : II :
DB 78 SACSGCQAKNWPPLAAQAQA 99

RESULT 10
Q9W2C6 PRELIMINARY; PRT; 2931 AA.
AC Q9W2C6;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2001 (TReMBLrel. 16, Last sequence update)
DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
GN G13492 PROTEIN.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cui Y., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,

```

```

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003455; AAF46766.2; -.
DR FlyBase; FBgn0034662; CG13492.
DR InterPro; IPR000173; GAP.DH.
DR PROSITE; PS00071; GAPDH; UNKNOWN_1.
SQ SEQUENCE 2931 AA; 316518 MW; B06C15D9DF257EF6 CRC64;

Query Match 37.1%; Score 50.5; DB 5; Length 2931;
Best Local Similarity 27.5%; Pred. No. 25;
Matches 11; Conservative 7; Mismatches 9; Indels 13; Gaps 1;

QY 2 SAITXOCKENLPSCSXOG-----LRAEENI 28
II : II : II : II : II : II :
DB 57 GAVIERGCDNLASVCTESGDNCYCRSHGCFNLQENL 96

RESULT 11
O45893 PRELIMINARY; PRT; 358 AA.
AC O45893;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE W06D12.4 PROTEIN.
GN W06D12.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Basham V.M.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z82073; CAB04925.1; -.
DR InterPro; IPR003002; 7TM_chemorecept_1.
DR InterPro; IPR00168; 7TM_nematode.
DR Pfam; PF01461; 7tm_4; 1.
SQ SEQUENCE 358 AA; 41363 MW; DA440A1E08B82323 CRC64;

Query Match 34.6%; Score 47; DB 5; Length 358;
Best Local Similarity 44.4%; Pred. No. 10;
Matches 12; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 2 SAITXOCKENLPSCSXOGLRAEENI 28
II III : II : II : II : II : II :
DB 227 GIATKTKIRKDLKSLRSLRLELNL 253

RESULT 12
O75136 PRELIMINARY; PRT; 1207 AA.
AC O75136;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE KIAA0641 PROTEIN.

```


OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RX MEDLINE=20083487; PubMed=10617197;

RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., VanAken S.E., Umavam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 RT thaliana.";
 RN [2]

RN Nature 402:761-768(1999).

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC004680; AAC31855.1; -

DR InterPro; IPR002086; Aldehyde_dehydr.

DR InterPro; IPR002114; PTS_HPr_ser.

DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.

DR PROSITE; PS00589; PTS_HPr_ser; UNKNOWN_1.

SQ SEQUENCE 1002 AA; 111924 MW; 225EC3EF7C8EC438 CRC64;

Query Match 33.8%; Score 46; DB 10; Length 1002;
 Best Local Similarity 60.0%; Pred. NO. 46;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 TGSATXQCKENLPS 15

Db 437 SGSMIDQCKSEKLP 451

RESULT 16

Q94BS5

ID Q94BS5 PRELIMINARY; PRT; 1002 AA.

AC Q94BS5;

DT 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE HYPOTHETICAL l11.9 KDA PROTEIN.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,

RA Dale J.M., Gibson H.A., Goldsmith A.D., Jiang P.X., Lee J.M.,

RA Quach H.L., Tang C., Toriumi M., Yu G., Bowser L., Carninci P.,

RA Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,

RA Karlin-Neumann G., Kawai J., Kim C., Koeseema E., Lam B., Lin J.,

RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,

RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E.,

RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;

RT "Full Length cDNA of gene F23F1.11/At2g29970 (GI:3420054).";

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

RX EMBL; AY039922; AAK64026.1; -

KW Hypothetical protein.

SQ SEQUENCE 1002 AA; 111866 MW; E15EC3104F8EC59A CRC64;

Query Match 33.8%; Score 46; DB 10; Length 1002;
 Best Local Similarity 60.0%; Pred. NO. 46;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 TGSATXQCKENLPS 15

Db 437 SGSMIDQCKSEKLP 451

RESULT 17

Q9DAT2

ID Q9DAT2 PRELIMINARY; PRT; 218 AA.

AC Q9DAT2;

DT 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)

DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)

DE 1600027N09RIK PROTEIN.

GN 1600027N09RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=PLACENTA;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Fesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

RA Wyszaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,

RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

DR EMBL; AK005551; BAB24117.1; -

DR MGD; MGI:1920497; 1600027N09RIK.

SQ SEQUENCE 218 AA; 23887 MW; A31C1C8895073463 CRC64;

Query Match 33.1%; Score 45; DB 11; Length 218;

Best Local Similarity 38.1%; Pred. NO. 13;

Matches 8; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

OY 1 TGSATXQCKENLPSLCXOG 21

Db 179 TGLSTRSDCNPNPACQCEG 199

RESULT 18

Q9CN47

ID Q9CN47 PRELIMINARY; PRT; 559 AA.

AC Q9CN47;

DT 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE HYPOTHETICAL PROTEIN PM0597.

GN PM0597.

OS Pasteurella multocida.

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

OC Pasteurella.

OX NCBI_TaxID=747;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PM70;

RX MEDLINE=21145866; PubMed=11248100;


```
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK018116; BAB31078.1; -.
DR HSSP: Q99418; 1PBV
DR MGD: MGI:1921252; 6330404E20R1k.
DR InterPro: IPR001849; PH.
DR InterPro: IPR000904; Sec7.
DR InterPro: IPR001605; Spectrin_PH.
DR Pfam: PF001169; PH; 1.
DR PRINTS: PR00683; SPECTRINPH.
DR SMART: SM00233; PH; 1.
DR PROSITE: PS50003; PH_DOMAIN; 1.
DR PROSITE: PS50190; SEC7; 1.
SQ SEQUENCE 360 AA; 40761 MW; 91B18E5210FA5275 CRC64;

Query Match 32.4%; Score 44; DB 11; Length 360;
Best Local Similarity 37.9%; Pred. No. 34;
Matches 11; Conservative 5; Mismatches 9; Indels 4; Gaps 1;

QY 3 SAITXCKENLPS----LCSXQGLRAEN 27
DB 224 SSMKKFCRPLLPSTRLCOEQLRSHEN 252

RESULT 22
QSDUT6 PRELIMINARY; PRT; 377 AA.
AC Q9DUT6;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE MEMBRANE PROTEIN.
GN M. Porcine rubulavirus.
OS Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Rubulavirus.
OX NCBI_TaxID=53179;
RN [1]
RP SEQUENCE FROM N.A.
RA Klenk C., Klenk H.D.;
RT "Sequence characterization of the membrane protein of porcine
parainfluenza virus (SER).";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ278914; CAC07420.1; -.
DR InterPro: IPR000982; Matrix.
DR Pfam: PF00661; Matrix; 1.
DR ProDom: PD000741; Matrix; 1.
SQ SEQUENCE 377 AA; 42123 MW; 8AB6FBF5AB586A49 CRC64;

Query Match 32.4%; Score 44; DB 12; Length 377;
Best Local Similarity 40.0%; Pred. No. 35;
Matches 10; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 TGSAITXCKENLPSLCSXQGLRAE 25
DB 85 TRAMVTLGCGPNLPSLGNVLGQLSE 109

RESULT 23
ID O13792 PRELIMINARY; PRT; 434 AA.
AC O13792; p78919;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PUTATIVE ATP-DEPENDENT RNA HELICASE C17G6.14C.
GN SPAC17G6.14C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
```

```
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RA Murphy L., Harris D., Barrell B.G., Rajandream M.A., Wood V.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 96-434 FROM N.A.
RC STRAIN-PR745;
RX MEDLINE=98162722; PubMed=9501991;
RA Yoshioka S., Kato K., Nakai K., Okayama H., Nojima H.;
RL DNA Res. 4:363-369(1997).
CC -!- SIMILARITY: TO OTHER "DEAD" BOX FAMILY HELICASES. BELONGS TO THE
CC DECD SUBFAMILY.
DR EMBL: Z99162; CAB16225.1; -.
DR EMBL: D89270; BAA13931.1; -.
DR HSSP: Q58083; 1HV8.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR001650; Helicase_C.
DR Pfam: PF00270; DEAD; 1.
DR Pfam: PF00271; helicase_C; 1.
DR SMART: SM00487; DEXDc; 1.
DR SMART: SM00490; HELICc; 1.
KW Hypothetical protein; ATP-binding; RNA-binding; Helicase.
FT DOMAIN 12 16 POLY-GLU.
FT NP_BIND 95 102 ATP (POTENTIAL).
FT SITE 204 207 DEAD BOX.
FT CONFLICT 192 192 I -> N (IN REF. 2).
SQ SEQUENCE 434 AA; 49231 MW; 7DBBC7EDE01FC3DB CRC64;

Query Match 32.4%; Score 44; DB 3; Length 434;
Best Local Similarity 52.9%; Pred. No. 41;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 12 NLPFSLCSXQGLRAEENI 28
DB 319 NFPSICHGGLPQEERI 335

RESULT 24
Q9PEZ3 PRELIMINARY; PRT; 443 AA.
AC Q9PEZ3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN XF0885.
GN XF0885.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OX Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorriy H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
```

RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meldandis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*.";
 RL Nature 406:151-159(2000).
 DR EMBL: AE003928; AAF83695.1; -.
 DR InterPro: IPR001296; Glycos_transf_1.1.
 DR Pfam: PF00534; Glycos_transf_1; 1.
 KW Hypothetical protein: Complete proteome.
 SQ SEQUENCE 443 AA; 49249 MW; 1CAEA51224C59A97 CRC64;

Query Match 32.4%; Score 44; DB 16; Length 443;
 Best Local Similarity 34.8%; Pred. No. 42;
 Matches 8; Conservative 3; Mismatches 12; Indels 0; Gaps 0;
 QY 2 GSAITXQCKENLPSCSXOGLRA 24
 I:: I:: I:: I:: I:: I::
 Db 114 GYKVVHLCQERMPSAEVALRA 136
 RESULT 25
 Q9BQ17 PRELIMINARY; PRT; 771 AA.
 AC Q9BQ17
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE HYPOTHETICAL 84.7 KDA PROTEIN.
 GN DKFZP761B0514.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE-AMYGDALA;
 RX MEDLINE=21154917; PubMed=11230166;
 RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
 RA Ansorge W., Boecker M., Bloecker H., Bauersachs S., Blum H.,
 RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,
 RA Mewes H.W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
 RA Wambutt R., Korn B., Klein M., Poustka A.;
 RT "Towards a Catalog of Human Genes and Proteins: Sequencing and
 RT Analysis of 500 Novel Complete Protein Coding Human cDNAs.";
 RL Genome Res. 11:422-435(2001).
 DR EMBL: AL136559; CAB66494.1; -.
 DR HSSP: Q99418; 1PBV.
 DR InterPro: IPR001849; PH.
 DR InterPro: IPR000904; Sec7.
 DR InterPro: IPR001605; Spectrin_PH.
 DR Pfam: PF00169; PH; 1.
 DR PRINTS: PR00683; SPECTRINPH.
 DR SMART: SM00233; PH; 1.
 DR SMART: SM00222; Sec7; 1.
 DR PROSITE: PS00003; PH_DOMAIN; 1.
 DR PROSITE: PS50190; SEC7; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 771 AA; 84660 MW; 3D7E08229429ED78 CRC64;

Query Match 32.4%; Score 44; DB 4; Length 771;
 Best Local Similarity 37.9%; Pred. No. 76;
 Matches 11; Conservative 5; Mismatches 9; Indels 4; Gaps 1;

QY 3 SAITXQCKENLPSCSXOGLRAEEN 27
 I:: I:: I:: I:: I:: I::
 Db 637 SSMKFCRPLLPSCYTRLCQEQELRSHEN 665
 RESULT 26
 Q9Y5D0 PRELIMINARY; PRT; 817 AA.
 ID Q9Y5D0
 AC Q9Y5D0
 DT 01-NOV-1999 (TRENBLrel. 12, Created)
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE PROTOCADHERIN GAMMA A7 SHORT FORM PROTEIN.
 GN PCDH-GAMMA-A7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99308636; PubMed=10380929;
 RA Wu Q., Maniatis T.;
 RT "A striking organization of a large family of human neural cadherin-
 RT like cell adhesion genes.";
 RL Cell 97:779-790(1999).
 CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.
 DR EMBL: AF152514; AAD43774.1; -.
 DR HSSP: P15116; INCJ.
 DR InterPro: IPR002126; Cadherin.
 DR Pfam: PF00028; cadherin; 6.
 DR PRINTS: PR00205; CADHERIN.
 DR SMART: SM00112; CA; 6.
 DR PROSITE: PS00232; CADHERIN_1; 1.
 DR PROSITE: PS50268; CADHERIN_2; 6.
 KW Calcium-binding; Cell adhesion; Glycoprotein.
 SQ SEQUENCE 817 AA; 89717 MW; 4B956F3AFB34F022 CRC64;

Query Match 32.4%; Score 44; DB 4; Length 817;
 Best Local Similarity 77.8%; Pred. No. 81;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 8 QCKENLPSL 16
 I:: I:: I:: I:: I:: I::
 Db 799 ECKENLPSCI 807
 RESULT 27
 Q9Y5G6 PRELIMINARY; PRT; 932 AA.
 ID Q9Y5G6
 AC Q9Y5G6
 DT 01-NOV-1999 (TRENBLrel. 12, Created)
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE PROTOCADHERIN GAMMA A7.
 GN PCDH-GAMMA-A7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RX MEDLINE=99308636; PubMed=10380929;
 RA Wu Q., Maniatis T.;
 RT "A striking organization of a large family of human neural cadherin-
 RT like cell adhesion genes.";
 RL Cell 97:779-790(1999).
 CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.
 DR EMBL: AF152327; AAD43721.1; -.
 DR HSSP: P15116; INCJ.
 DR InterPro: IPR002126; Cadherin.

```
DR Pfam; PF00028; cadherin; 6.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 6.
DR PROSITE; PS00232; CADHERIN.1; 1.
DR PROSITE; PS00268; CADHERIN.2; 6.
KW Calcium-binding; Cell adhesion; Glycoprotein.
SQ SEQUENCE 932 AA; 101721 MW; 9614B69C0CA2371 CRC64;

Query Match          32.4%; Score 44; DB 4; Length 932;
Best Local Similarity 77.8%; Pred. No. 93;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 8 QCKENLPSL 16
   :|||||:
Db 799 ECKENLPST 807

RESULT 28
Q962J8          PRELIMINARY; PRT; 1011 AA.
ID Q962J8;
AC Q962J8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PV1H14180_P.
GN PV1H14180W.
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RA Tchavtchitch M., Fischer K., Huestis R., Saul A.;
RT "The sequence of 200 kb portion of a plasmodium vivax chromosome
RT reveals a high degree of conservation with P. falciparum chromosome
RT 3."
RL Mol. Biochem. Parasitol. 0:0-0(2001).
DR EMBL; AF003872; AAF99482.1; -.
SQ SEQUENCE 1011 AA; 113202 MW; BDE5D72F8573BC80 CRC64;

Query Match          32.4%; Score 44; DB 5; Length 1011;
Best Local Similarity 32.0%; Pred. No. 1e+02;
Matches 8; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 3 SAITXQCKENLPSCSXQGLRAEEN 27
   | : | | | | : | : | | | : | : |
Db 694 STLTGCKKALEGQICREKNLREQQN 718

RESULT 29
O73554          PRELIMINARY; PRT; 243 AA.
ID O73554;
AC O73554;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 27.5 KDA PROTEIN.
GN FGF.
OS Spodoptera littoralis nuclear polyhedrosis virus (SLNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=10456;
RN [1]
RP SEQUENCE FROM N.A.
RA Martin O., Smith I., Croizier L., Lopez-Ferber M., Durantel D.,
RA Hunter-Fujita F., Crolzier G.;
RT "Comparative analysis of open reading frames in the egt gene regions
RT of Spodoptera littoralis and S. litura nucleopolyhedroviruses."
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ003131; CAA05888.1; -.
DR HSSP; P31371; 1G82.
DR InterPro; IPR002209; HBGF_FGF.

DR Pfam; PF00167; FGF; 1.
DR ProDom; PD000831; HBGF_FGF; 1.
DR SMART; SM00442; FGF; 1.
KW Hypothetical protein.
SQ SEQUENCE 243 AA; 27451 MW; B6EB329494514F60 CRC64;

Query Match          32.0%; Score 43.5; DB 12; Length 243;
Best Local Similarity 34.4%; Pred. No. 27;
Matches 11; Conservative 6; Mismatches 8; Indels 7; Gaps 2;

QY 2 GSAITXQCKEN-----LPS--LCSXQGLRAEE 26
   | : | | : | | : | | : | : |
Db 163 GPSITNRCIENDKRNIMPSSGVCEGDGAGED 194

RESULT 30
O59353          PRELIMINARY; PRT; 111 AA.
ID O59353;
AC O59353;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DE 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE HYPOTHETICAL 11.6 KDA PROTEIN PH1649.
GN PH1649.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oouchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
RL DNA Res. 5:55-76(1998).
DR EMBL; AP000006; BAA30761.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 111 AA; 11590 MW; A607A025243FA8A1 CRC64;

Query Match          31.6%; Score 43; DB 17; Length 111;
Best Local Similarity 38.1%; Pred. No. 14;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 3 SAITXQCKENLPSCSXQGLR 23
   | : | : | | | | | | | | : |
Db 13 SPLKLRQCEKLPINSSQGIK 33

RESULT 31
P97337          PRELIMINARY; PRT; 146 AA.
ID P97337;
AC P97337;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ODORANT BINDING PROTEIN IB (FRAGMENT).
GN OBP1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=NASAL MUCOSA;
RX MEDLINE=98326309; PubMed=9661663;
RA Pes D., Mameli M., Andreini I., Krelger J., Weber M., Breers H.,
```



```
OC Spermatoophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Paillard M., Sederoff R.R., Levings C.S. III;
RT "Nucleotide sequence of the S-1 mitochondrial Dna from the cytoplasm
of maize.";
RL EMBO J. 4:1125-1128(1985).
DR EMBL; X02451; CAB57810.1; -.
KW Mitochondrion.
FT NON_TER
SQ SEQUENCE 256 AA; 29308 MW; DAFA69F8039CA893 CRC64;

Query Match 31.6%; Score 43; DB 8; Length 256;
Best Local Similarity 44.0%; Pred. No. 35;
Matches 11; Conservative 2; Mismatches 8; Indels 4; Gaps 1;

OY 2 GSAITXCKENLPSCSXOGLRAEE 26
DB 199 GSA----CREGVDSCKEKGCGEE 219

RESULT 35
ID Q98SV1 PRELIMINARY; PRT; 367 AA.
AC Q98SV1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SELENOPROTEIN PA.
OS Brachydanio rerio (Zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21110713; PubMed=11168591;
RX Kryukov G.V., Gladyshev V.N.;
RT "Selenium metabolism in zebrafish: multiplicity of selenoprotein genes
and expression of a protein containing 17 selenocysteine residues.";
RL Genes Cells 5:1049-1060(2000).
DR EMBL; AF322071; AAG53688.1; -.
SQ SEQUENCE 367 AA; 41289 MW; A0BA8D94A09FFA41 CRC64;

Query Match 31.6%; Score 43; DB 13; Length 367;
Best Local Similarity 47.4%; Pred. No. 51;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 9 CKENLPSCSXOGLRAEN 27
DB 312 CDEPLPASXPQGLKEQDN 330

RESULT 36
ID Q9LFN6 PRELIMINARY; PRT; 427 AA.
AC Q9LFN6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE DEAD BOX RNA HELICASE RH15.
GN F2111_90.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatoophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RA Bevan M., Peters S.A., van Staveren M., Dirkse W., Stiekema W.,
RA Bancroft I., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL360314; CAB96655.1; -.
DR HSSP; Q58083; 1HV8.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICc; 1.
KW ATP-binding; Helicase.
SQ SEQUENCE 427 AA; 48384 MW; B673D11A9B6D1271 CRC64;

Query Match 31.6%; Score 43; DB 10; Length 427;
Best Local Similarity 41.2%; Pred. No. 60;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 12 NLPSLCSXOGLRAEENI 28
DB 312 NFPSICIHGMSQEERL 328

RESULT 37
ID Q93VJ8 PRELIMINARY; PRT; 427 AA.
AC Q93VJ8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE AT5G11200/F2111_90 (AT5G11170/F2111_60).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatoophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis ORF clones.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Koesema E., Meyers M.C.,
RA Banh J., Bowser L., Carninci P., Chung M.K., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yu G., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cdna clones.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
```


Search completed: August 22, 2002, 07:52:19
Job time: 197 sec

2002-08-22 08:07:24

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 11:16:53 ; Search time 27.33 Seconds
(without alignments)
101.961 Million cell updates/sec

Title: US-09-856-199-1
Perfect score: 29
Sequence: 1 KRCINQLLCKLPDSELPASXCCHCRHL 29

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1200 summaries

Database : PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	72.4	381	1 A47327	selenoprotein p pr
2	9	31.0	380	2 T10442	selenoprotein p pr
3	9	31.0	385	1 OMRTSP	selenoprotein p pr
4	7	24.1	258	2 AF2287	group 3 sigma 37-t
5	7	24.1	559	2 A25872	transcription regu
6	7	24.1	1009	2 T47327	hypothetical prote
7	6	20.7	129	2 D70026	hypothetical prote
8	6	20.7	133	2 C24925	lactose permease -
9	6	20.7	180	1 RKMUA1	ribulose-bisphosph
10	6	20.7	180	2 G56694	hypothetical prote
11	6	20.7	205	2 A71177	probable thymidyla
12	6	20.7	207	2 T36656	probable carbonic
13	6	20.7	217	2 A10870	hypothetical prote
14	6	20.7	228	2 T20219	hypothetical prote
15	6	20.7	236	2 E50674	hypothetical prote
16	6	20.7	236	2 H85524	hypothetical prote
17	6	20.7	259	2 B95267	probable ABC trans
18	6	20.7	281	2 A82104	conserved hypothet
19	6	20.7	303	2 D70955	hypothetical prote
20	6	20.7	310	1 D64758	probable transcrip
21	6	20.7	321	2 AB0317	lipic acid synthe
22	6	20.7	346	2 F70666	probable alcohol d
23	6	20.7	348	2 AF2505	hypothetical prote
24	6	20.7	359	2 T16350	hypothetical prote
25	6	20.7	364	2 C75597	thymidine diphosph
26	6	20.7	388	2 T54525	leukemia-related p
27	6	20.7	395	1 S36186	mitogen-activated
28	6	20.7	416	2 JT0487	lactose permease -
29	6	20.7	461	2 G85059	probable sugar tra

30	6	20.7	478	2 JK0201	alpha-amylase (EC
31	6	20.7	498	2 A48305	alpha-amylase (EC
32	6	20.7	499	1 ALAS1	alpha-amylase (EC
33	6	20.7	499	1 ALAS3	alpha-amylase (EC
34	6	20.7	499	2 B48305	alpha-amylase (EC
35	6	20.7	499	2 JS0663	alpha-amylase (EC
36	6	20.7	499	2 JT0466	alpha-amylase (EC
37	6	20.7	499	2 JN0588	alpha-amylase (EC
38	6	20.7	504	2 AD1615	glycine betaine/ca
39	6	20.7	504	2 AF1252	glycine betaine/ca
40	6	20.7	505	2 AB0528	deoxyguanosinetrip
41	6	20.7	609	2 T22803	hypothetical prote
42	6	20.7	622	2 T22716	hypothetical prote
43	6	20.7	683	2 T24486	hypothetical prote
44	6	20.7	776	2 T09938	hypothetical prote
45	6	20.7	929	2 A44048	genome polyprotein
46	6	20.7	999	1 S27756	receptor-like prot
47	6	20.7	1102	2 S28104	probable DNA-direc
48	6	20.7	1151	2 S03722	DNA-directed DNA p
49	6	20.7	1319	2 S75705	hypothetical prote
50	6	20.7	2301	1 GNNYTN	genome polyprotein
51	6	20.7	2303	1 GNNYTM	genome polyprotein
52	6	20.7	2303	1 GNNYTP	genome polyprotein
53	6	20.7	2303	2 S13554	hypothetical prote
54	6	20.7	2824	2 T22759	hypothetical prote
55	5	17.2	29	2 I52402	alpha-fetoprotein
56	5	17.2	34	2 A44038	leader peptide Tna
57	5	17.2	44	2 A60329	antigen PV9 - Plas
58	5	17.2	53	2 S22326	gene K protein - p
59	5	17.2	53	2 H82738	hypothetical prote
60	5	17.2	54	2 G82682	hypothetical prote
61	5	17.2	60	2 S44215	genome polyprotein
62	5	17.2	76	2 AG0685	hypothetical prote
63	5	17.2	79	2 G87186	conserved hypothet
64	5	17.2	83	2 D87646	preprotein transio
65	5	17.2	83	2 G84620	hypothetical prote
66	5	17.2	84	1 Q6BP57	gene 6.5 protein -
67	5	17.2	85	2 AE2483	hypothetical prote
68	5	17.2	86	2 AD0356	probable [4Fe-4S]
69	5	17.2	87	2 G70782	hypothetical prote
70	5	17.2	88	2 C83823	hypothetical prote
71	5	17.2	94	2 T03310	hypothetical prote
72	5	17.2	95	2 AC1903	hypothetical prote
73	5	17.2	98	1 BORT2	prostatic steroid-
74	5	17.2	99	2 G83295	hypothetical prote
75	5	17.2	103	2 S56123	alpha-amylase/subt
76	5	17.2	107	2 A70129	ribosomal protein
77	5	17.2	110	2 B70565	hypothetical prote
78	5	17.2	111	2 T14304	ribosomal protein
79	5	17.2	112	2 T45914	60S RIBOSOMAL PROT
80	5	17.2	113	2 T02526	60S RIBOSOMAL PROT
81	5	17.2	117	2 C87588	hypothetical prote
82	5	17.2	119	2 S61276	gene flhD protein
83	5	17.2	119	2 T49715	hypothetical prote
84	5	17.2	121	2 T29680	hypothetical prote
85	5	17.2	123	2 S56967	hypothetical prote
86	5	17.2	123	2 A64206	hypothetical prote
87	5	17.2	123	2 T14741	hypothetical prote
88	5	17.2	125	2 T11941	ribosomal protein
89	5	17.2	125	2 A35003	galactose-binding
90	5	17.2	126	2 F82072	holo-(acyl-carrier
91	5	17.2	128	2 A99559	hypothetical prote
92	5	17.2	129	2 S72304	hypothetical prote
93	5	17.2	131	2 S78133	ribosomal protein
94	5	17.2	132	2 B83379	hypothetical prote
95	5	17.2	135	2 S76237	hypothetical prote
96	5	17.2	138	2 T39484	hypothetical prote
97	5	17.2	139	2 G75127	hypothetical prote
98	5	17.2	139	2 E71003	hypothetical prote
99	5	17.2	139	2 G44032	hypothetical prote
100	5	17.2	140	2 AC1264	hypothetical prote
101	5	17.2	140	2 AE1626	hypothetical prote
102	5	17.2	142	1 F70437	nucleoside-diphosp

103	5	17.2	148	2	C75212	hypothetical prote	176	5	17.2	226	2	AC1072	conserved hypothet
104	5	17.2	150	2	S47459	probable membrane	177	5	17.2	227	2	T29531	hypothetical prote
105	5	17.2	151	2	C72555	hypothetical prote	178	5	17.2	230	2	T37016	probable oxidoredu
106	5	17.2	152	2	E95018	conserved domain p	179	5	17.2	230	2	D81252	hypothetical prote
107	5	17.2	152	2	F97891	hypothetical prote	180	5	17.2	230	2	AF2237	hypothetical prote
108	5	17.2	153	2	G97221	uncharacterized co	181	5	17.2	230	2	C83181	hypothetical prote
109	5	17.2	157	2	AI2451	hypothetical prote	182	5	17.2	231	2	G87442	ribonuclease III f
110	5	17.2	158	2	D72016	hypothetical prote	183	5	17.2	231	2	AB1202	E. coli copper hom
111	5	17.2	158	2	H86606	hypothetical prote	184	5	17.2	231	2	AH1559	E. coli copper hom
112	5	17.2	158	2	AE2511	hypothetical prote	185	5	17.2	235	2	D70948	probable pilli asse
113	5	17.2	159	2	C83604	conserved hypothet	186	5	17.2	237	2	C83380	DNA polymerase III
114	5	17.2	159	2	T37098	hypothetical prote	187	5	17.2	237	2	E87249	hypothetical prote
115	5	17.2	160	2	T29392	hypothetical prote	188	5	17.2	238	2	SI2259	ferritin precursor
116	5	17.2	161	1	FIAAA	translational initia	189	5	17.2	239	2	S45604	hypothetical prote
117	5	17.2	161	2	D87406	ribosomal protein	190	5	17.2	240	2	D96637	hypothetical prote
118	5	17.2	164	2	T06913	H ⁺ -transporting AT	191	5	17.2	241	2	AF0956	phosphate transpor
119	5	17.2	164	2	B72717	hypothetical prote	192	5	17.2	241	2	F71478	probable pseudouri
120	5	17.2	165	2	H65049	ygad protein - Esc	193	5	17.2	241	2	T17798	hypothetical prote
121	5	17.2	165	2	E91073	hypothetical prote	194	5	17.2	241	2	S74620	hypothetical prote
122	5	17.2	165	2	A85918	hypothetical prote	195	5	17.2	242	2	D40590	transcription init
123	5	17.2	165	2	AI0843	conserved hypothet	196	5	17.2	243	2	F90721	probable glutamate
124	5	17.2	167	2	T09686	TCTP protein homol	197	5	17.2	243	2	D85572	probable chaperone
125	5	17.2	167	2	F64998	hypothetical prote	198	5	17.2	243	2	S42533	uroporphyrinogen I
126	5	17.2	167	2	G91023	hypothetical prote	199	5	17.2	243	2	D71684	hypothetical prote
127	5	17.2	167	2	H85867	hypothetical prote	200	5	17.2	243	2	C97748	hypothetical prote
128	5	17.2	168	1	C64888	conserved hypothet	201	5	17.2	244	2	S40698	riboflavin biosynt
129	5	17.2	168	2	E90878	probable filament	202	5	17.2	246	2	AB0319	probable glutamate
130	5	17.2	168	2	T08831	disease resistance	203	5	17.2	246	2	A25779	retrovirus-related
131	5	17.2	168	2	E85740	hypothetical prote	204	5	17.2	245	2	AH0926	hypothetical prote
132	5	17.2	169	2	E83578	disulfide bond for	205	5	17.2	247	2	AG1300	3-ketoacyl-acyl ca
133	5	17.2	169	2	T08824	disease resistance	206	5	17.2	247	2	AG1672	3-ketoacyl-acyl ca
134	5	17.2	171	2	B97652	probable acetyltra	207	5	17.2	248	1	S43421	ribosomal protein
135	5	17.2	171	2	AI2875	acetyltransferase f	208	5	17.2	248	1	H70447	3-oxoacyl-facyl-ca
136	5	17.2	174	2	A70395	hypothetical prote	209	5	17.2	249	1	A45351	immediate-early pr
137	5	17.2	176	2	F71540	hypothetical prote	210	5	17.2	249	2	H98130	probable short-cha
138	5	17.2	176	2	AB0777	probable lipoprote	211	5	17.2	249	2	AI3156	short-chain dehydr
139	5	17.2	179	2	T03043	hypothetical prote	212	5	17.2	249	2	T38475	hypothetical prote
140	5	17.2	180	2	C83573	hypothetical prote	213	5	17.2	250	2	E84321	50S ribosomal prot
141	5	17.2	181	2	E82333	oligoribonuclease	214	5	17.2	253	2	F87371	transcription regu
142	5	17.2	185	2	F72528	hypothetical prote	215	5	17.2	254	2	F82484	conserved hypothet
143	5	17.2	186	2	PW0008	chaperonine 60K be	216	5	17.2	255	2	AF0491	probable fimbrial
144	5	17.2	188	2	F97554	hypothetical prote	217	5	17.2	256	2	G82170	conserved hypothet
145	5	17.2	193	2	T19081	hypothetical prote	218	5	17.2	256	2	H82235	response regulator
146	5	17.2	196	2	G90115	hypothetical prote	219	5	17.2	258	2	H75616	oxidoreductase, sh
147	5	17.2	198	2	C97641	teillurite resistan	220	5	17.2	258	2	T46491	hypothetical prote
148	5	17.2	198	2	AD2864	methyltransferase	221	5	17.2	259	2	T44454	arginine/ornithine
149	5	17.2	200	1	A37146	ribosomal protein	222	5	17.2	262	2	S57276	14-3-3 protein hom
150	5	17.2	201	2	T44489	hypothetical prote	223	5	17.2	262	2	F90895	hypothetical prote
151	5	17.2	201	2	H81449	probable dechibio	224	5	17.2	262	2	H85722	hypothetical prote
152	5	17.2	203	2	S04860	alpha-amylase/subt	225	5	17.2	262	2	T31238	trAK protein homol
153	5	17.2	203	2	S67198	hypothetical prote	226	5	17.2	264	2	H85090	14-3-3-like protei
154	5	17.2	204	2	S22639	alpha-amylase/subt	227	5	17.2	267	2	H84396	hypothetical prote
155	5	17.2	206	2	E82427	pyrazinamidase/nic	228	5	17.2	268	2	T46980	hypothetical prote
156	5	17.2	208	2	G96724	hypothetical prote	229	5	17.2	268	2	AE0241	conserved hypothet
157	5	17.2	208	2	FQ0278	hypothetical prote	230	5	17.2	269	2	B44499	major intrinsic pr
158	5	17.2	209	2	D84120	uracil phosphoribo	231	5	17.2	269	2	A41616	erythrocyte integr
159	5	17.2	210	2	S52050	cytochrome-c oxida	232	5	17.2	269	2	JC1320	water channel prot
160	5	17.2	210	2	AG0950	ATP/GTP-binding pr	233	5	17.2	269	2	I52366	uterine water chan
161	5	17.2	210	2	T22284	hypothetical prote	234	5	17.2	271	2	JC2348	water channel prot
162	5	17.2	213	2	D82917	adenylate kinase U	235	5	17.2	273	2	F97712	hypothetical prote
163	5	17.2	214	2	JC5043	granulocyte colony	236	5	17.2	273	2	B70550	probable ciste prot
164	5	17.2	215	2	A61516	variant surface gl	237	5	17.2	273	2	G87037	conserved hypothet
165	5	17.2	218	2	D81516	conserved hypothet	238	5	17.2	274	2	H84956	2,3,4,5-tetrahydro
166	5	17.2	220	2	T36936	hypothetical prote	239	5	17.2	274	2	C97647	hypothetical prote
167	5	17.2	221	2	F83048	probable two-compo	240	5	17.2	274	2	AI2870	metal dependent hy
168	5	17.2	224	2	S23728	MADS box protein T	241	5	17.2	275	2	E95231	pur operon repress
169	5	17.2	225	1	C64245	peripheral membran	242	5	17.2	275	2	G98095	activator of purin
170	5	17.2	225	2	S56598	probable phosphata	243	5	17.2	277	2	G69158	conserved hypothet
171	5	17.2	225	2	I38700	hNKR-P1a protein -	244	5	17.2	277	2	F31619	A44L protein - var
172	5	17.2	225	2	D91295	probable phosphata	245	5	17.2	277	2	F36852	hypothetical prote
173	5	17.2	225	2	C97342	transcription regu	246	5	17.2	277	2	T28583	12L protein - vari
174	5	17.2	225	2	F86136	probable phosphata	247	5	17.2	277	2	D42521	A38L protein - vac
175	5	17.2	226	2	JQ1263	hypothetical 26K p	248	5	17.2	277	2	S29922	SalL8L protein pre

249	5	17.2	277	2	G83908	hypothetical prote	322	5	17.2	324	2	S48122	transcription fact
250	5	17.2	277	2	T51327	L-aspartate oxidas	323	5	17.2	326	2	T26647	hypothetical prote
251	5	17.2	277	2	S57381	protein disulfide	324	5	17.2	327	2	A56545	F1b1 protein - yea
252	5	17.2	277	2	AE3296	glutamine amidotra	325	5	17.2	328	2	F81317	UDPglucose 4-epime
253	5	17.2	278	2	AF2636	methyltransferase	326	5	17.2	329	2	AI2023	prochlorophyllid
254	5	17.2	280	2	D70803	hypothetical prote	327	5	17.2	329	2	AD1698	diaminopimelate ep
255	5	17.2	280	2	H69288	conserved hypotet	328	5	17.2	330	2	T50135	probable l-lactate
256	5	17.2	280	2	T41986	glycoprotein - hum	329	5	17.2	330	2	T46256	brevican - human (
257	5	17.2	280	2	AC1238	hypothetical prote	330	5	17.2	331	2	AE2928	dehydrogenase Atu3
258	5	17.2	282	2	AC3231	hypothetical prote	331	5	17.2	331	2	B70336	heterodisulfide re
259	5	17.2	283	2	T31147	hypothetical prote	332	5	17.2	332	2	F83600	conserved hypotet
260	5	17.2	283	2	B98113	hypothetical prote	333	5	17.2	333	2	H82245	glyceraldehyde 3-p
261	5	17.2	283	2	AH1767	protoporphyrinogen	334	5	17.2	333	2	I38974	G protein-coupled
262	5	17.2	285	1	QOECHF	hydrogenase-1 cofa	335	5	17.2	334	2	T18983	hypothetical prote
263	5	17.2	285	2	E90770	hypothetical prote	336	5	17.2	334	2	AC3282	primosomal protein
264	5	17.2	285	2	A85633	hypothetical prote	337	5	17.2	335	2	GH1280	holliday junction
265	5	17.2	286	2	A86480	F14D7.8 protein -	338	5	17.2	335	2	G83598	cell division prot
266	5	17.2	287	2	S71279	L-ascorbate peroxi	339	5	17.2	336	2	G82228	forminoglutamase
267	5	17.2	287	2	E86196	hypothetical prote	340	5	17.2	336	2	AE0761	conserved hypotet
268	5	17.2	288	2	T09845	L-ascorbate peroxi	341	5	17.2	337	2	I48682	NEX-1 - mouse
269	5	17.2	289	2	G86858	conserved hypotet	342	5	17.2	337	2	I57038	gene Dlx-3 protein
270	5	17.2	289	2	C96778	hypothetical prote	343	5	17.2	338	2	T03432	probable transport
271	5	17.2	290	2	F69897	transcription regu	344	5	17.2	338	2	B82698	glucanotactonase p
272	5	17.2	290	2	T08692	hypothetical prote	345	5	17.2	338	2	C87311	conserved hypotet
273	5	17.2	291	2	F75172	hypothetical prote	346	5	17.2	339	2	G90760	probable integrase
274	5	17.2	292	2	D86859	LysR family transc	347	5	17.2	339	2	E82211	conserved hypotet
275	5	17.2	292	2	H64129	hemK protein (EC 1	348	5	17.2	339	2	S37920	MBR1 protein precu
276	5	17.2	294	1	S73832	MG237 homolog F10_	349	5	17.2	339	2	E85624	probable integrase
277	5	17.2	296	2	E82250	conserved hypotet	350	5	17.2	339	2	AC3245	hypothetical prote
278	5	17.2	296	2	F84633	probable carbonyl	351	5	17.2	341	2	JC7149	cell division prot
279	5	17.2	296	2	T47914	hypothetical prote	352	5	17.2	341	2	T41450	hypothetical prote
280	5	17.2	297	2	T22690	hypothetical prote	353	5	17.2	342	1	G70460	conserved hypotet
281	5	17.2	297	2	S55085	hypothetical prote	354	5	17.2	342	2	H71258	protein Tp33 precu
282	5	17.2	297	2	H69708	DNA processing Smf	355	5	17.2	342	2	E98302	luxa-related prote
283	5	17.2	297	2	S17474	hypothetical prote	356	5	17.2	342	2	AB2981	monooxygenase [imp
284	5	17.2	299	2	C86196	hypothetical prote	357	5	17.2	343	2	D84469	potassium channel
285	5	17.2	299	2	B86196	hypothetical prote	358	5	17.2	343	2	T22018	hypothetical prote
286	5	17.2	299	2	A54030	PxR protein - Chin	359	5	17.2	343	2	AI0581	DNA polymerase III
287	5	17.2	299	2	I37468	house keeping gene	360	5	17.2	344	2	A98354	hypothetical prote
288	5	17.2	299	2	H90748	probable surface p	361	5	17.2	344	2	T00993	probable beta-1,3-
289	5	17.2	299	2	D85599	probable surface p	362	5	17.2	344	2	T33057	hypothetical prote
290	5	17.2	300	2	E97886	hypothetical prote	363	5	17.2	345	2	T29786	hypothetical prote
291	5	17.2	301	2	G96013	conserved hypotet	364	5	17.2	346	2	S72475	glucanase [EC 3.2.
292	5	17.2	301	2	C82791	conserved hypotet	365	5	17.2	346	2	AB2260	glucanase [importe
293	5	17.2	302	1	T1RTGK	tissue factor path	366	5	17.2	346	2	T32273	hypothetical prote
294	5	17.2	302	2	G75206	maltoextrin trans	367	5	17.2	348	2	T50783	protein phosphatas
295	5	17.2	303	2	G82652	cytochrome C oxida	368	5	17.2	349	2	G86813	hypothetical prote
296	5	17.2	303	2	C37994	RF3 protein - saim	369	5	17.2	350	2	E75341	peptidyl-prolyl ci
297	5	17.2	304	2	G87291	integrase/recombin	370	5	17.2	351	2	JQ2327	ALL protein - Indi
298	5	17.2	305	2	F84982	hypothetical prote	371	5	17.2	351	2	AI3355	UDP-3-O-[3-hydroxy
299	5	17.2	307	2	B87388	sirohaem synthase	372	5	17.2	351	2	D96761	unknown protein [i
300	5	17.2	308	2	T25508	hypothetical prote	373	5	17.2	352	2	D70788	probable trbB prot
301	5	17.2	308	2	S32172	neurosporene hydro	374	5	17.2	353	2	T11275	NADH dehydrogenase
302	5	17.2	309	2	AB0493	probable formate d	375	5	17.2	353	2	AF2053	hypothetical prote
303	5	17.2	311	2	B26955	hypothetical prote	376	5	17.2	354	2	AD0374	phosphoribosylamin
304	5	17.2	313	2	F64433	hypothetical prote	377	5	17.2	355	1	DCECPK	phosphoribosylamin
305	5	17.2	314	2	H82384	tagE protein VCA10	378	5	17.2	355	2	H90701	phosphoribosylamin
306	5	17.2	314	2	E97418	hypothetical prote	379	5	17.2	355	2	C85552	phosphoribosylamin
307	5	17.2	315	2	B64826	probable membrane	380	5	17.2	355	2	AG0568	phosphoribosylamin
308	5	17.2	315	2	C71344	probable spore coa	381	5	17.2	358	1	A48952	triacylglycerol li
309	5	17.2	316	2	D70773	probable thrB prot	382	5	17.2	358	2	A64424	hydrogenase expres
310	5	17.2	318	2	H82493	MoxR-related prote	383	5	17.2	358	2	D82485	transcription regu
311	5	17.2	318	2	C71173	hypothetical prote	384	5	17.2	359	2	S22593	hypothetical prote
312	5	17.2	319	2	B75495	conserved hypotet	385	5	17.2	359	2	S39211	gene Cl protein -
313	5	17.2	319	2	T46594	phytoene synthase	386	5	17.2	361	2	H84949	peptide chain rele
314	5	17.2	319	2	T23194	hypothetical prote	387	5	17.2	362	1	H64132	phosphoribosylamin
315	5	17.2	320	2	F75082	hypothetical prote	388	5	17.2	362	2	S12933	killer toxin K2 pr
316	5	17.2	320	2	T25702	hypothetical prote	389	5	17.2	363	2	JT0599	triacylglycerol li
317	5	17.2	321	1	S27501	alkaline proteinas	390	5	17.2	364	1	A39133	triacylglycerol li
318	5	17.2	321	2	F70141	probable oligopept	391	5	17.2	364	1	WEPGA	inhibin alpha chai
319	5	17.2	323	2	S28671	noCR protein - Agr	392	5	17.2	364	2	C97393	hypothetical prote
320	5	17.2	324	2	S35462	NADH dehydrogenase	393	5	17.2	364	2	AE2611	conserved hypotet
321	5	17.2	324	2	B96742	NAM-like protein,	394	5	17.2	365	2	S47657	peroxidase (EC 1.1

395	5	17.2	365	2	JC4726	manganese peroxidase	468	5	17.2	405	2	A72383	sensor histidine kinase
396	5	17.2	365	2	S75244	potassium channel	469	5	17.2	405	2	A82130	transcription regulator
397	5	17.2	369	2	D82098	3,4-dihydroxy-2-butanone	470	5	17.2	407	2	AG2415	two-component hybrid
398	5	17.2	370	2	A95416	probable enoyl reductase	471	5	17.2	409	2	T47470	isovaleryl-CoA dehydrogenase
399	5	17.2	370	2	F87638	conserved hypothetical protein	472	5	17.2	409	2	F83493	probable MFS transporter
400	5	17.2	370	2	A49528	probable structural protein	473	5	17.2	410	2	AD2844	conserved hypothetical protein
401	5	17.2	371	2	D84186	hypothetical protein	474	5	17.2	411	2	G82159	interleukin-1 receptor
402	5	17.2	371	2	B64907	hypothetical protein	475	5	17.2	411	2	B72102	conserved hypothetical protein
403	5	17.2	372	2	T00243	sopa protein - Esc	476	5	17.2	411	2	E86519	CT178 hypothetical protein
404	5	17.2	372	2	T32370	hypothetical protein	477	5	17.2	411	2	D86995	probable D-alanyl-tRNA synthetase
405	5	17.2	373	2	AC0253	ribonuclease III (RNase D)	478	5	17.2	411	2	F87644	transcription regulator
406	5	17.2	373	2	F97621	hypothetical protein	479	5	17.2	413	2	E70177	acetate kinase (ACK1)
407	5	17.2	373	2	T40602	hypothetical protein	480	5	17.2	414	2	S06729	probable IS117 transposase
408	5	17.2	374	2	AD1668	phosphoribosylamine transferase	481	5	17.2	415	2	AG2044	hypothetical protein
409	5	17.2	374	2	AF1296	phosphoribosylamine transferase	482	5	17.2	416	2	B83745	aspartate transaminase
410	5	17.2	374	2	D81715	conserved hypothetical protein	483	5	17.2	416	2	B64132	tRNA adenylyltransferase
411	5	17.2	374	2	E89920	conserved hypothetical protein	484	5	17.2	416	2	T20300	hypothetical protein
412	5	17.2	375	1	NRECD	ribonuclease III (RNase D)	485	5	17.2	416	2	S75793	poly(A) polymerase
413	5	17.2	375	2	A99943	RNase D, processes	486	5	17.2	416	2	S33473	interleukin-1 receptor
414	5	17.2	375	2	E85791	ribonuclease D (RNase D)	487	5	17.2	420	1	EDBE17	immediate-early protein
415	5	17.2	375	2	AB0725	hypothetical protein	488	5	17.2	420	2	S25236	85K surface antigen
416	5	17.2	376	2	S45107	hypothetical protein	489	5	17.2	421	2	A60058	neural cell adhesion molecule
417	5	17.2	376	2	C95988	probable oxidoreductase	490	5	17.2	421	2	T48202	protein kinase AK2
418	5	17.2	376	2	AE0599	probable inner membrane protein	491	5	17.2	421	2	T40845	hypothetical protein
419	5	17.2	377	2	AC1723	hypothetical protein	492	5	17.2	423	2	A37033	isovaleryl-CoA dehydrogenase
420	5	17.2	377	2	AH1352	hypothetical protein	493	5	17.2	423	2	B86657	hypothetical protein
421	5	17.2	377	2	G90737	hypothetical protein	494	5	17.2	424	2	B38176	samB protein - Sal
422	5	17.2	377	2	A64816	ybhS protein - Esc	495	5	17.2	424	2	AB1034	UV protection protein
423	5	17.2	377	2	H85587	hypothetical protein	496	5	17.2	424	2	B89015	protein B0213.11 (Hsp70)
424	5	17.2	378	1	DEECPP	probable erythronucleoside	497	5	17.2	426	2	G82309	probable hemolysin
425	5	17.2	378	2	A32916	long-chain-fatty-acyl-CoA synthetase	498	5	17.2	428	2	E87633	adenylosuccinate synthetase
426	5	17.2	378	2	D91029	erythronate-4-phosphate	499	5	17.2	428	2	A83642	hypothetical protein
427	5	17.2	378	2	E85873	erythronate-4-phosphate	500	5	17.2	430	2	A87541	major facilitator
428	5	17.2	378	2	AE0802	erythronate-4-phosphate	501	5	17.2	430	2	AD1611	probable proteinase
429	5	17.2	379	1	DCBSPK	phosphoribosylamine transferase	502	5	17.2	430	2	AH1248	probable proteinase
430	5	17.2	379	2	H83727	phosphoribosylamine transferase	503	5	17.2	431	2	S15471	sol protein - Ther
431	5	17.2	379	2	T24054	hypothetical protein	504	5	17.2	433	2	A72445	probable hemolysin
432	5	17.2	379	2	C90987	hypothetical protein	505	5	17.2	433	2	D75632	melibiase (alpha-galactosidase)
433	5	17.2	379	2	E64972	probable polysaccharide	506	5	17.2	434	2	D83928	hypothetical protein
434	5	17.2	379	2	F85832	probable polysaccharide	507	5	17.2	434	2	G98235	hypothetical protein
435	5	17.2	379	2	AC0770	probable polysaccharide	508	5	17.2	435	2	T36780	hypothetical protein
436	5	17.2	379	2	F83152	hypothetical protein	509	5	17.2	435	2	T35884	tagatose 6-phosphate
437	5	17.2	381	2	S34848	homocitrate synthase	510	5	17.2	436	2	JC4742	transposase - Cory
438	5	17.2	381	2	S37297	nif protein - Kle	511	5	17.2	437	2	A75194	glycogen synthase
439	5	17.2	382	1	B46233	transcription factor	512	5	17.2	437	2	B86510	KDO transferase [1]
440	5	17.2	383	2	S76146	hypothetical protein	513	5	17.2	437	2	E81557	3-deoxy-manno-8-phosphate
441	5	17.2	383	2	F69018	formate dehydrogenase	514	5	17.2	437	2	S70546	3-deoxy-manno-8-phosphate
442	5	17.2	383	2	T05590	hypothetical protein	515	5	17.2	437	2	S55392	msh protein - fruit
443	5	17.2	383	2	E75210	hypothetical protein	516	5	17.2	438	2	T50629	hypothetical protein
444	5	17.2	384	2	T32756	hypothetical protein	517	5	17.2	439	2	B81701	phospholipase D family
445	5	17.2	384	2	T32756	hypothetical protein	518	5	17.2	441	1	JN0482	3-phosphate (EC 3.1.1.1)
446	5	17.2	386	2	AD1273	N-acetylornithine	519	5	17.2	442	2	AG2296	hypothetical protein
447	5	17.2	386	2	T28176	hypothetical protein	520	5	17.2	444	2	C83336	probable cytochrome
448	5	17.2	387	2	S60222	probable chlorophyll	521	5	17.2	444	2	T47114	probable 3-carboxy
449	5	17.2	387	2	F84392	hypothetical protein	522	5	17.2	445	2	I49458	acetylcholine receptor
450	5	17.2	388	1	BVECAF	sopa protein - Esc	523	5	17.2	445	2	AE3050	two component sensor
451	5	17.2	388	2	JU0053	NADPH-dependent butyrate	524	5	17.2	449	2	B71265	probable Mg2+ transporter
452	5	17.2	388	2	S76932	probable phosphoribosyl	525	5	17.2	450	2	T50192	probable pseudouridine
453	5	17.2	389	2	F97692	glutamate 5-kinase	526	5	17.2	451	2	G82218	sensor histidine kinase
454	5	17.2	389	2	AC2918	glutamate 5-kinase	527	5	17.2	452	2	A82540	outer membrane export
455	5	17.2	389	2	A97495	hypothetical protein	528	5	17.2	452	2	A10779	probable n-hydroxy
456	5	17.2	389	2	AC2713	conserved hypothetical protein	529	5	17.2	453	2	T37054	hypothetical protein
457	5	17.2	390	1	A39249	neutrophil cytosolic factor	530	5	17.2	453	2	A89873	hypothetical protein
458	5	17.2	394	2	A75303	UDP-glucose 4-epimerase	531	5	17.2	456	1	ACCHAN	nicotinic acetylcholine
459	5	17.2	395	1	B49335	acetate kinase (EC 2.7.1.3)	532	5	17.2	456	2	S75034	thiophen / furan oxygenase
460	5	17.2	395	2	S70912	CMP-N-acetylneuraminase	533	5	17.2	457	1	ACHUA1	nicotinic acetylcholine
461	5	17.2	396	2	G48327	COI intron 6 protein	534	5	17.2	457	1	ACBOA1	nicotinic acetylcholine
462	5	17.2	397	2	AE1272	acetate kinase homologue	535	5	17.2	457	2	A69359	amidophosphoribosyl
463	5	17.2	397	2	AG1634	acetate kinase homologue	536	5	17.2	457	2	S08162	nicotinic acetylcholine
464	5	17.2	400	2	C89955	hypothetical protein	537	5	17.2	457	2	S13872	nicotinic acetylcholine
465	5	17.2	400	2	T32705	hypothetical protein	538	5	17.2	457	2	A24383	nicotinic acetylcholine
466	5	17.2	400	2	T25605	hypothetical protein	539	5	17.2	457	2	A28529	nicotinic acetylcholine
467	5	17.2	403	2	T28416	ORF MSV255 leucine	540	5	17.2	457	2	E70070	metabolite transporter

541	5	17.2	457	2	S76703	hypothetical prote	614	5	17.2	513	2	AE1921	hypothetical prote
542	5	17.2	460	2	A12264	aldehyde dehydrog	615	5	17.2	514	2	T43304	Rad17 protein homo
543	5	17.2	461	1	ACRYAL	nicotinic acetylch	616	5	17.2	515	2	G97396	probable chelatase
544	5	17.2	461	2	I50548	acetylcholine rece	617	5	17.2	516	2	G98198	proline/betaine tr
545	5	17.2	461	2	H64087	seryl-tRNA(Sec) se	618	5	17.2	516	2	AB3088	MFS permease [prol
546	5	17.2	462	2	JC5625	14-nm filament pro	619	5	17.2	519	2	B70603	probable methionyl
547	5	17.2	462	2	T28699	hypothetical prote	620	5	17.2	520	2	B71143	hypothetical prote
548	5	17.2	462	2	T17480	endo-xy lanase homo	621	5	17.2	522	2	B48956	4-chlorobenzoate--
549	5	17.2	463	1	G59602	late competence pr	622	5	17.2	523	2	F86481	59.2K hypothetical
550	5	17.2	463	2	E72268	astCB/chuk-related	623	5	17.2	525	2	T00459	hypothetical prote
551	5	17.2	463	2	T51194	hypothetical prote	624	5	17.2	525	2	S70528	sensor protein pil
552	5	17.2	464	2	C64462	adenosylmethionine	625	5	17.2	526	2	E89015	protein B0213.15 l
553	5	17.2	465	1	S47738	cytochrome-c perox	626	5	17.2	526	2	E75045	2-isopropylmalate
554	5	17.2	465	2	T03560	probable aldehyde	627	5	17.2	526	2	B67272	3',5'-cyclic-nucle
555	5	17.2	465	2	F91178	cytochrome-c perox	628	5	17.2	528	2	H70140	oligopeptide ABC t
556	5	17.2	465	2	G86024	cytochrome-c perox	629	5	17.2	528	2	B75364	extracellular solu
557	5	17.2	467	1	JN0889	3-phytase (EC 3.1.	630	5	17.2	528	2	T50330	hypothetical prote
558	5	17.2	467	1	JN0656	3-phytase (EC 3.1.	631	5	17.2	528	2	D86456	probable protein k
559	5	17.2	467	1	UFEC	fumurate hydratase	632	5	17.2	531	2	AG1812	hypothetical prote
560	5	17.2	467	2	E90918	fumurate C [import	633	5	17.2	532	1	RRVOC2	RNA-directed RNA p
561	5	17.2	467	2	B85767	fumarase C [import	634	5	17.2	532	2	T04748	hypothetical prote
562	5	17.2	469	2	T09366	cytochrome P450 ho	635	5	17.2	534	2	S31300	regulatory protein
563	5	17.2	470	2	B64039	hypothetical prote	636	5	17.2	535	2	C87266	comm protein [limpo
564	5	17.2	470	2	A87298	hypothetical prote	637	5	17.2	537	2	G31277	guinate transport
565	5	17.2	472	2	T08860	hypothetical prote	638	5	17.2	537	2	F86938	probable methionyl
566	5	17.2	473	2	B84556	hypothetical prote	639	5	17.2	538	2	T46863	L-aspartate oxidas
567	5	17.2	474	2	T27297	hypothetical prote	640	5	17.2	538	2	A40640	pentachlorophenol-
568	5	17.2	475	1	CZCLCA	cellulase (EC 3.2.	641	5	17.2	538	2	S76085	hypothetical prote
569	5	17.2	475	2	AG0268	indole-3-glycerol-	642	5	17.2	540	1	OXECLD	L-aspartate oxidas
570	5	17.2	476	2	G90416	hypothetical prote	643	5	17.2	540	2	H91058	quinolinolate synth
571	5	17.2	476	2	D86412	protein FIK23.6 [i	644	5	17.2	540	2	F85903	quinolinolate synth
572	5	17.2	478	2	A54248	heparin cofactor I	645	5	17.2	540	2	AB0830	L-aspartate oxidas
573	5	17.2	478	2	D83043	conserved hypothet	646	5	17.2	542	2	F90170	conserved hypothet
574	5	17.2	479	1	S41066	heparin cofactor I	647	5	17.2	545	1	JX0225	cytochrome P450 Cy
575	5	17.2	479	2	G86207	hypothetical prote	648	5	17.2	545	2	AF0330	L-aspartate oxidas
576	5	17.2	479	2	T15065	hypothetical prote	649	5	17.2	545	2	T43513	Hsk1 protein kinas
577	5	17.2	480	2	T24422	amidophosphoribos	650	5	17.2	550	2	B82074	1-aspartate oxidas
578	5	17.2	480	2	C84967	pyruvate kinase (E	651	5	17.2	550	2	T19257	hypothetical prote
579	5	17.2	480	2	I46990	heparin cofactor I	652	5	17.2	551	2	G98665	hypothetical prote
580	5	17.2	481	2	T49321	related to heterok	653	5	17.2	553	2	S76757	hypothetical prote
581	5	17.2	484	2	H96505	probable aldehyde	654	5	17.2	553	2	T22370	hypothetical prote
582	5	17.2	486	2	C81332	hypothetical prote	655	5	17.2	564	2	E82539	hypothetical prote
583	5	17.2	487	2	E70961	variant surface gl	656	5	17.2	567	2	H84770	ABC transporter AR
584	5	17.2	487	2	E90251	aspartate oxidase	657	5	17.2	568	2	S17996	probable receptor-
585	5	17.2	489	2	S23410	FUN19 protein - ve	658	5	17.2	568	2	T31692	hypothetical prote
586	5	17.2	493	2	S77612	amidophosphoribos	659	5	17.2	569	2	E86957	probable cholesterol
587	5	17.2	493	2	S68587	nicotinic acetylch	660	5	17.2	570	2	T32743	hypothetical prote
588	5	17.2	493	2	T24695	hypothetical prote	661	5	17.2	571	1	XZADH5	penton protein (II
589	5	17.2	494	2	S15523	serine proteinase,	662	5	17.2	571	1	C70353	succinate dehydrog
590	5	17.2	497	2	B81728	probable do serine	663	5	17.2	575	2	S77416	succinate dehydrog
591	5	17.2	497	2	H71465	probable secreted	664	5	17.2	575	2	AC2177	succinate dehydrog
592	5	17.2	497	2	AF0506	protein B0213.12 [665	5	17.2	575	2	A97635	adenine deaminase
593	5	17.2	499	2	C89015	protein B0213.14 [666	5	17.2	575	2	AC2858	hypothetical prote
594	5	17.2	499	2	D89015	protein B0213.16 [667	5	17.2	575	2	B83959	G protein-coupled
595	5	17.2	499	2	F89015	complement C1 inhi	668	5	17.2	576	2	A48765	L-aspartate oxidas
596	5	17.2	500	1	ITHUC1	hypothetical prote	669	5	17.2	578	2	AF1958	hypothetical prote
597	5	17.2	500	2	T33279	betaine-aldehyde d	670	5	17.2	579	1	D70740	fibrin-like 71 K
598	5	17.2	500	2	T41385	hypothetical prote	671	5	17.2	579	2	JC7170	L-aspartate oxidas
599	5	17.2	500	2	T16630	hypothetical prote	672	5	17.2	580	2	T36393	hypothetical prote
600	5	17.2	500	2	T28168	hemolysin secretio	673	5	17.2	584	2	B76210	cholesterol oxidas
601	5	17.2	500	2	AB2517	probable ABC subst	674	5	17.2	585	2	S72824	hypothetical prote
602	5	17.2	501	2	G82896	L-aspartate oxidas	675	5	17.2	586	2	D84710	probable adenylate
603	5	17.2	502	2	E81780	cytochrome P450 ho	676	5	17.2	587	2	E95369	chaperonin 62.5K b
604	5	17.2	502	2	H81203	probable exonuclea	677	5	17.2	588	2	PW0007	hypothetical prote
605	5	17.2	506	2	T04628	I2 protein - Trypa	678	5	17.2	591	2	B71920	protein secretion
606	5	17.2	506	2	F96772	conserved hypothet	679	5	17.2	591	2	S64727	developmental prot
607	5	17.2	506	2	S47439	L-aspartate oxidas	680	5	17.2	595	2	T06412	probable chaperoni
608	5	17.2	507	2	H69186	Bifunctional phosph	681	5	17.2	597	2	A30098	development-specif
609	5	17.2	509	2	D84339	hypothetical prote	682	5	17.2	599	2	S18612	probable chaperoni
610	5	17.2	509	2	AE1295	conserved hypothet	683	5	17.2	600	2	JT0901	chaperonin 60 beta
611	5	17.2	509	2	A84404	L-aspartate oxidas	684	5	17.2				
612	5	17.2	510	2	AG2614		685	5	17.2				
613	5	17.2	511	2	G87609		686	5	17.2				

587	5	17.2	600	2	B96597	Rubisco subunit bi	760	5	17.2	683	1	RN2MB1	DNA-directed RNA p
588	5	17.2	600	2	T06292	hypothetical prote	761	5	17.2	683	1	D83481	probable oligopept
589	5	17.2	603	2	D81906	probable GTP-bind	762	5	17.2	684	2	T41595	probable zinc fing
590	5	17.2	603	2	H81111	Gnp-binding protei	763	5	17.2	685	2	AC0527	ferrichrome transp
591	5	17.2	604	2	A11959	two-component sens	764	5	17.2	687	2	T04793	probable fimbrin A
592	5	17.2	606	2	A54665	netrin-1 precursor	765	5	17.2	691	2	T46476	hypothetical prote
593	5	17.2	606	2	A11855	hypothetical prote	766	5	17.2	698	2	T32840	hypothetical prote
594	5	17.2	608	2	C87106	probable penicilli	767	5	17.2	699	2	T14904	NADPH--ferrhemopr
595	5	17.2	609	2	T52524	hypothetical prote	768	5	17.2	699	2	S25541	heat shock protein
596	5	17.2	611	1	FPRT	alpha-fetoprotein	769	5	17.2	699	2	T07037	heat shock protein
597	5	17.2	611	2	C81388	succinate dehydrog	770	5	17.2	699	2	A96802	unknown protein [i
598	5	17.2	614	2	B82396	translation initia	771	5	17.2	700	2	T24092	hypothetical prote
599	5	17.2	615	2	T34727	transketolase - St	772	5	17.2	702	1	JQ0868	glucan 1,4-alpha-g
700	5	17.2	616	2	H84474	probable Athila re	773	5	17.2	705	2	A45508	heat shock protein
701	5	17.2	618	2	T49177	hypothetical prote	774	5	17.2	706	2	T15701	hypothetical prote
702	5	17.2	619	2	D90072	hypothetical prote	775	5	17.2	706	2	AF2494	hypothetical prote
703	5	17.2	621	2	A71961	90Kda chaperone -	776	5	17.2	707	2	T40070	origin recognition
704	5	17.2	621	2	B64546	chaperone and heat	777	5	17.2	709	2	C87258	fatty oxidation co
705	5	17.2	622	2	D82055	conserved hypothet	778	5	17.2	712	2	T16338	hypothetical prote
706	5	17.2	627	2	B71709	hypothetical prote	779	5	17.2	715	2	A48426	heat shock protein
707	5	17.2	630	2	AC1309	probable ABC trans	780	5	17.2	716	2	G96538	hypothetical prote
708	5	17.2	630	2	AC1681	probable ABC trans	781	5	17.2	719	1	S70583	replication licens
709	5	17.2	630	2	D95937	probable mureinpep	782	5	17.2	719	1	JC4580	replication licens
710	5	17.2	630	2	D97669	hypothetical prote	783	5	17.2	719	2	F96830	hypothetical prote
711	5	17.2	630	2	A12893	hypothetical prote	784	5	17.2	719	2	A30047	enhancer of split
712	5	17.2	633	2	T47654	hypothetical prote	785	5	17.2	719	2	AC0867	2-acylglycerophosp
713	5	17.2	634	1	B32392	protein kinase C (786	5	17.2	720	2	T47221	replication licens
714	5	17.2	634	1	GERTX1	matrix glycoprotei	787	5	17.2	721	2	D82934	ATP-dependent zinc
715	5	17.2	634	2	AD2343	hypothetical prote	788	5	17.2	724	1	WMBEP	infected cell prot
716	5	17.2	634	2	T22351	hypothetical prote	789	5	17.2	726	2	D86390	TIK7.25 protein -
717	5	17.2	635	2	T06537	hypothetical prote	790	5	17.2	730	1	T03920	replication licens
718	5	17.2	637	2	T08050	probable ethylene	791	5	17.2	733	2	S31288	MAK10 protein - ye
719	5	17.2	638	2	H82690	hypothetical prote	792	5	17.2	734	2	S46765	hypothetical prote
720	5	17.2	640	2	T51231	H-protein promoter	793	5	17.2	734	2	T24908	hypothetical prote
721	5	17.2	640	2	A41726	homeotic protein B	794	5	17.2	736	2	T40904	hypothetical prote
722	5	17.2	640	2	T25367	hypothetical prote	795	5	17.2	737	2	C81724	pombe specific hyp
723	5	17.2	642	2	T51421	L-aspartate oxidas	796	5	17.2	743	2	T38674	1,4-alpha-glucan b
724	5	17.2	642	2	S58154	hypothetical prote	797	5	17.2	743	2	T32421	probable membrane
725	5	17.2	642	2	T12113	transcription fact	798	5	17.2	744	2	A45622	hypothetical prote
726	5	17.2	642	2	S27806	homeotic protein B	799	5	17.2	744	2	F75625	serine proteinase,
727	5	17.2	644	2	T16136	hypothetical prote	800	5	17.2	745	2	B83834	hypothetical prote
728	5	17.2	645	2	C64979	molR_2 protein - E	801	5	17.2	747	2	T42599	minor capsid prote
729	5	17.2	648	2	T40504	atp-dependent, dea	802	5	17.2	747	2	T47385	hypothetical prote
730	5	17.2	650	2	H81708	hypothetical prote	803	5	17.2	750	1	ALLVP7	photosystem I P700
731	5	17.2	652	2	A84461	probable fimbrin [804	5	17.2	753	1	WZBEB8	gene 56 protein -
732	5	17.2	653	2	B82872	transketolase I UU	805	5	17.2	753	2	T07543	photosystem I prot
733	5	17.2	653	2	F70864	probable oxidoredu	806	5	17.2	753	2	A96747	probable RNA-bind
734	5	17.2	655	2	T30044	hypothetical prote	807	5	17.2	753	2	S48059	metal-regulatory t
735	5	17.2	659	1	B44212	structural protein	808	5	17.2	757	2	G71981	probable transcrip
736	5	17.2	659	2	A64139	ribonuclease T2 ho	809	5	17.2	757	2	T48417	hypothetical prote
737	5	17.2	660	2	A64739	ferrichrome-iron t	810	5	17.2	760	2	AB2225	hypothetical prote
738	5	17.2	660	2	B90648	hypothetical prote	811	5	17.2	761	2	T03719	probable thyroide r
739	5	17.2	660	2	B85499	hypothetical prote	812	5	17.2	768	2	H82367	sensory box/GGDF
740	5	17.2	662	2	S58298	ATPase - pepper (f	813	5	17.2	769	2	H64525	transcription regu
741	5	17.2	662	2	S42826	probable ATPase -	814	5	17.2	771	2	T16254	hypothetical prote
742	5	17.2	663	2	S67259	MNE1 protein - Yea	815	5	17.2	773	2	C84554	hypothetical prote
743	5	17.2	664	2	S60062	hevin precursor -	816	5	17.2	775	2	B72074	hypothetical prote
744	5	17.2	664	2	T50316	hypothetical Armad	817	5	17.2	775	2	C81594	hypothetical prote
745	5	17.2	667	1	E64240	hypothetical prote	818	5	17.2	775	2	D86549	hypothetical prote
746	5	17.2	668	2	T15305	hypothetical prote	819	5	17.2	779	2	T05990	hypothetical prote
747	5	17.2	669	2	S46518	luciferin-binding	820	5	17.2	783	2	F88808	protein R09E10.3 [
748	5	17.2	669	2	S46519	luciferin-binding	821	5	17.2	784	2	T45027	hypothetical prote
749	5	17.2	670	1	A30882	arachidonate 5-lip	822	5	17.2	786	2	S67060	probable membrane
750	5	17.2	670	1	S52478	carnitine O-acetyl	823	5	17.2	793	2	JC5539	smoothed protein
751	5	17.2	670	2	S52637	phosphoenolpyruvat	824	5	17.2	795	2	F75154	cell division cont
752	5	17.2	673	2	S46520	luciferin-binding	825	5	17.2	798	2	B71196	probable transio
753	5	17.2	673	2	S35335	transcription fact	826	5	17.2	798	2	S62791	probable lipoprote
754	5	17.2	674	2	T49479	arachidonate 5-lip	827	5	17.2	802	1	A37142	outer membrane ush
755	5	17.2	675	2	T04644	hypothetical prote	828	5	17.2	805	2	S13410	chloride channel p
756	5	17.2	675	2	S25786	hypothetical prote	829	5	17.2	807	2	T00990	hypothetical prote
757	5	17.2	676	1	A39379	hatching-suppres	830	5	17.2	809	2	S16266	cellulose synthase
758	5	17.2	676	2	T23690	hypothetical prote	831	5	17.2	810	2	S19725	voltage-gated chor
759	5	17.2	682	1	RNR2C1	DNA-directed RNA p	832	5	17.2	811	2	AH2552	hypothetical prote

833	5	17.2	812	2	E95029	formate acetyltran	906	5	17.2	1036	2	B69368	hypothetical prote
834	5	17.2	815	2	H97900	formate C-acetyltr	907	5	17.2	1039	2	G86203	probable N-arginin
835	5	17.2	822	2	T47007	hypothetical prote	908	5	17.2	1046	2	T30199	chitinase (EC 3.2.
836	5	17.2	822	2	A80238	heamin storage sys	909	5	17.2	1051	2	D82428	chitodextrinase VC
837	5	17.2	822	2	S56823	probable membrane	910	5	17.2	1054	2	T30933	chitinase (EC 3.2.
838	5	17.2	826	2	G90283	hypothetical prote	911	5	17.2	1061	2	H90084	hypothetical prote
839	5	17.2	838	2	T04785	hypothetical prote	912	5	17.2	1073	1	S56822	SK12 protein homol
840	5	17.2	840	2	H71114	probable cell divi	913	5	17.2	1085	2	S40476	Ca(2+)-sensing rec
841	5	17.2	840	2	D75046	transitional endop	914	5	17.2	1088	1	PIXRBR	inner layer protei
842	5	17.2	840	2	S41218	H1R1 protein - yea	915	5	17.2	1088	1	PIXRBR	inner layer protei
843	5	17.2	841	2	D84513	probable retroelem	916	5	17.2	1088	2	S39261	VPI protein - porc
844	5	17.2	843	2	T12689	armadillo segment	917	5	17.2	1088	2	S13558	VPI protein - bovi
845	5	17.2	857	2	S01864	lipoxigenase (EC 1	918	5	17.2	1088	2	H84604	probable cellulose
846	5	17.2	857	2	S44883	lipoxigenase (EC 1	919	5	17.2	1091	2	B95103	hypothetical prote
847	5	17.2	858	2	T12142	lipoxigenase (EC 1	920	5	17.2	1091	2	A58532	glial cell membran
848	5	17.2	860	2	AB2044	adenylate cyclase	921	5	17.2	1097	1	RNEVL3	DNA-directed DNA p
849	5	17.2	864	2	S13381	lipoxigenase (EC 1	922	5	17.2	1099	2	H83210	probable phospholi
850	5	17.2	868	2	T06827	lipoxigenase (EC 1	923	5	17.2	1102	2	T12681	hypothetical prote
851	5	17.2	868	2	T02635	D2 protein homolog	924	5	17.2	1108	2	D96798	hypothetical prote
852	5	17.2	883	2	S49126	brevican precursor	925	5	17.2	1113	2	T14260	period protein Per
853	5	17.2	883	2	S57653	brevican precursor	926	5	17.2	1115	1	IJMSNL	neural cell adhesi
854	5	17.2	884	2	C70729	hypothetical prote	927	5	17.2	1115	2	T13955	period protein Per
855	5	17.2	885	2	D90000	sensor protein kdp	928	5	17.2	1116	2	H97970	type I site-specif
856	5	17.2	888	2	B87270	tonB-dependent rec	929	5	17.2	1132	2	T00259	hypothetical prote
857	5	17.2	900	2	T32827	hypothetical prote	930	5	17.2	1132	2	T40355	hypothetical prote
858	5	17.2	906	2	B96901	uncharacterized co	931	5	17.2	1139	2	T33368	hypothetical prote
859	5	17.2	912	2	A54423	brevican precursor	932	5	17.2	1143	2	T28129	hypothetical prote
860	5	17.2	912	2	T02892	hypothetical prote	933	5	17.2	1150	2	S49956	probable membrane
861	5	17.2	915	2	AC3477	DNA mismatch repai	934	5	17.2	1153	2	T31080	nitric-oxide synth
862	5	17.2	915	2	S24577	ovarian protein -	935	5	17.2	1153	2	T21386	hypothetical prote
863	5	17.2	915	2	S74283	probable protein k	936	5	17.2	1159	2	T02866	hypothetical prote
864	5	17.2	916	2	T20909	hypothetical prote	937	5	17.2	1164	2	G71827	hypothetical prote
865	5	17.2	917	2	T50979	related to tol pro	938	5	17.2	1168	2	I56985	kalinin B1 - mouse
866	5	17.2	918	2	A89221	protein r27F2.2 li	939	5	17.2	1170	2	A53612	laminin B1k chain
867	5	17.2	921	2	F84593	hypothetical prote	940	5	17.2	1171	2	T17454	diaphanous-related
868	5	17.2	924	2	T21738	hypothetical prote	941	5	17.2	1179	2	T04584	TMV resistance pro
869	5	17.2	926	2	E83375	probable glycosyl	942	5	17.2	1197	1	G65010	sensor protein evg
870	5	17.2	931	2	G69615	ATP-dependent DNA	943	5	17.2	1197	2	A91035	probable sensor fo
871	5	17.2	932	2	T48761	hypothetical prote	944	5	17.2	1197	2	C85879	probable sensor fo
872	5	17.2	934	2	T47546	protein kinase-lik	945	5	17.2	1215	2	T00364	hypothetical prote
873	5	17.2	940	2	T01834	hypothetical prote	946	5	17.2	1217	2	T13996	pol protein - frul
874	5	17.2	944	1	S48821	probable membrane	947	5	17.2	1222	2	C88504	protein B0361.3 li
875	5	17.2	946	1	IYH02	inter-alpha-trypsi	948	5	17.2	1227	2	AG2504	hypothetical prote
876	5	17.2	946	2	S54354	inter-alpha-inhibi	949	5	17.2	1237	2	T08608	hypothetical prote
877	5	17.2	946	2	JC5575	inter-alpha-trypsi	950	5	17.2	1242	1	DJBEC1	DNA-directed DNA p
878	5	17.2	958	2	AC0204	probable integral	951	5	17.2	1247	2	T18671	hypothetical prote
879	5	17.2	959	2	T03053	ribonucleoside-dip	952	5	17.2	1251	2	B66194	hypothetical prote
880	5	17.2	962	2	C71617	SERA antigen/papai	953	5	17.2	1252	2	T14272	cortactin-binding
881	5	17.2	966	2	T30017	hypothetical prote	954	5	17.2	1254	1	A32686	DNA-directed DNA p
882	5	17.2	973	2	J00971	transposase tnpA -	955	5	17.2	1256	2	S60461	gene flightless-I
883	5	17.2	975	2	A86258	protein F5011.4 li	956	5	17.2	1262	2	T30524	protein phosphatas
884	5	17.2	981	2	H82403	chemotactic transd	957	5	17.2	1262	2	T13353	protein stn-B - fr
885	5	17.2	988	2	S77211	isoleucine - tRNA l	958	5	17.2	1271	2	T24008	hypothetical prote
886	5	17.2	988	2	A24341	transposase - Pseu	959	5	17.2	1274	2	A89959	hypothetical prote
887	5	17.2	988	2	JQ1477	transposase - Esch	960	5	17.2	1296	2	E81840	transcription repa
888	5	17.2	988	4	S58114	transposase (EC 6.	961	5	17.2	1302	2	T20767	hypothetical prote
889	5	17.2	989	2	C83035	hypothetical prote	962	5	17.2	1310	2	T40135	probable involveme
890	5	17.2	989	2	AE2140	toxin secretion AB	963	5	17.2	1316	2	T50444	hypothetical prote
891	5	17.2	991	2	T01372	hypothetical prote	964	5	17.2	1337	2	T38949	hypothetical prote
892	5	17.2	991	2	A99604	hypothetical prote	965	5	17.2	1346	2	G71613	hypothetical prote
893	5	17.2	999	2	S68689	glucose regulated	966	5	17.2	1379	2	A81102	transcription repa
894	5	17.2	1000	2	D87244	conserved hypothet	967	5	17.2	1385	2	T13887	tlr protein - frul
895	5	17.2	1003	2	C71139	hypothetical prote	968	5	17.2	1389	2	T13852	gene wheeler prote
896	5	17.2	1004	2	T30338	oviductin (EC 3.4.	969	5	17.2	1390	2	T31353	polyprotein - Arab
897	5	17.2	1020	2	JN0124	glycine dehydrogen	970	5	17.2	1390	2	T14004	trif protein - sil
898	5	17.2	1025	2	T44802	type I site-specif	971	5	17.2	1426	2	T30817	homeotic protein C
899	5	17.2	1025	2	T30322	type I site-specif	972	5	17.2	1460	2	T00095	hypothetical prote
900	5	17.2	1025	2	T09459	type I site-specif	973	5	17.2	1482	2	B43274	N-methyl-D-asparta
901	5	17.2	1026	2	B96663	hypothetical prote	974	5	17.2	1482	2	I49704	glutamate receptor
902	5	17.2	1026	2	T34506	hypothetical prote	975	5	17.2	1484	2	S22086	N-methyl-D-asparta
903	5	17.2	1029	2	AC3363	acriflavin resista	976	5	17.2	1496	2	A04047	insecticidal toxin
904	5	17.2	1031	2	H81288	probable sugar tra	977	5	17.2	1530	2	AD1663	glutamate synthase
905	5	17.2	1036	2	A82357	probable multidrug	978	5	17.2	1530	2	AF1291	glutamate synthase

979	5	17.2	1538	2	E70874	probable ppsB prot	1052	4	13.8	35	2	T07509	photosystem II pro
980	5	17.2	1548	2	S54723	unp-glucose--glyco	1053	4	13.8	35	1	S41911	hypothetical prote
981	5	17.2	1570	2	T38792	probable protein t	1054	4	13.8	37	1	R5IT36	ribosomal protein
982	5	17.2	1596	2	A41216	guanine nucleotide	1055	4	13.8	37	2	C83780	hypothetical prote
983	5	17.2	1630	2	T40217	hypothetical prote	1056	4	13.8	39	2	B61581	secretogranin II -
984	5	17.2	1679	2	S49802	probable membrane	1057	4	13.8	39	2	A47752	RNA recognition mo
985	5	17.2	1684	2	JW0057	gravin - human	1058	4	13.8	39	2	S63530	hypothetical prote
986	5	17.2	1706	2	B75633	probable RNA helic	1059	4	13.8	40	2	A95001	hypothetical prote
987	5	17.2	1713	2	A55347	adhesive ligand ep	1060	4	13.8	40	2	A47753	beta-defensin-10 -
988	5	17.2	1748	1	JN0786	integrin beta-4 ch	1061	4	13.8	42	2	D81005	hypothetical prote
989	5	17.2	1767	2	T20766	hypothetical prote	1062	4	13.8	43	2	S44463	defensin - Pyrrhoc
990	5	17.2	1770	2	S45842	tya protein - yeas	1063	4	13.8	43	2	A81532	hypothetical prote
991	5	17.2	1770	2	S69948	tya protein - yeas	1064	4	13.8	44	2	B87372	hypothetical prote
992	5	17.2	1770	2	A71517	hypothetical prote	1065	4	13.8	44	2	E85849	unknown protein en
993	5	17.2	1777	2	AC2088	serine/threonine k	1066	4	13.8	45	2	S71132	hypothetical prote
994	5	17.2	1780	2	A85045	probable glucan sy	1067	4	13.8	47	2	S10862	hypothetical prote
995	5	17.2	1791	2	T24089	hypothetical prote	1068	4	13.8	48	2	S32448	pol polyprotein -
996	5	17.2	1807	2	JC6319	integrin beta-4 ch	1069	4	13.8	49	2	S66455	taurochenodeoxycho
997	5	17.2	1808	2	AB1847	serine/threonine k	1070	4	13.8	49	2	D97489	hypothetical prote
998	5	17.2	1811	2	T39252	probable protein t	1071	4	13.8	50	1	INKS	insulin - casiragu
999	5	17.2	1847	2	T18308	probable vitellogen	1072	4	13.8	50	2	C69094	ribosomal protein
1000	5	17.2	1882	2	T00069	hypothetical prote	1073	4	13.8	50	2	H82032	hypothetical prote
1001	5	17.2	1932	2	AF1946	two-component hydr	1074	4	13.8	51	2	A56785	calmodulin - pig (
1002	5	17.2	1946	2	AE1449	hypothetical prote	1075	4	13.8	51	2	B90140	hypothetical prote
1003	5	17.2	1974	2	T16703	hypothetical prote	1076	4	13.8	51	2	AE2564	hypothetical prote
1004	5	17.2	2016	2	A38195	sodium channel pro	1077	4	13.8	52	2	A05273	glycophorin - dog
1005	5	17.2	2019	2	A33996	sodium channel pro	1078	4	13.8	52	2	H84150	hypothetical prote
1006	5	17.2	2049	2	T29227	hypothetical prote	1079	4	13.8	53	1	R3MX14	ribosomal protein
1007	5	17.2	2051	2	T13164	plexin B - fruit f	1080	4	13.8	53	1	LNLDA	lectin alpha chain
1008	5	17.2	2101	2	S57245	insulin receptor (1081	4	13.8	53	2	T59604	glutamate receptor
1009	5	17.2	2110	2	H96803	unknown protein T5	1082	4	13.8	53	2	A25989	mannose/glucose-sp
1010	5	17.2	2114	2	E96505	hypothetical prote	1083	4	13.8	53	2	JC6526	antibiotic mutaci
1011	5	17.2	2116	2	T49818	glutamate synthase	1084	4	13.8	53	2	G90776	hypothetical prote
1012	5	17.2	2140	2	T18543	probable cell-adhe	1085	4	13.8	54	2	H69960	exodeoxyribonuclea
1013	5	17.2	2148	1	A56081	insulin receptor -	1086	4	13.8	54	2	T48475	hypothetical prote
1014	5	17.2	2152	2	T45583	hypothetical prote	1087	4	13.8	54	2	C82603	hypothetical prote
1015	5	17.2	2248	1	D42088	adenylate cyclase	1088	4	13.8	54	2	A97784	hypothetical prote
1016	5	17.2	2367	2	S70172	toxin B - Clostrid	1089	4	13.8	55	2	F56976	transfer complex p
1017	5	17.2	2424	2	T46480	calcium channel BI	1090	4	13.8	55	2	T10344	hypothetical prote
1018	5	17.2	2493	2	T40540	hypothetical prote	1091	4	13.8	57	1	T1EPVA	venom basic protei
1019	5	17.2	2493	2	A55481	adenylate cyclase	1092	4	13.8	57	2	A61009	phosphopyruvate hy
1020	5	17.2	2688	2	T49477	alpha-A-crystallin	1093	4	13.8	57	2	S12957	venom animal Kunit
1021	5	17.2	2787	2	S45416	TEI1 protein - yea	1094	4	13.8	57	2	S63680	signal transducer
1022	5	17.2	2895	2	T08437	hyperplastic discs	1095	4	13.8	57	2	PC4415	Erbb kinase activa
1023	5	17.2	3131	2	S39842	enniatin synthetas	1096	4	13.8	57	2	B59399	short epsilon-dend
1024	5	17.2	3191	2	T22945	hypothetical prote	1097	4	13.8	58	2	A86747	glutaredoxin-like
1025	5	17.2	3216	2	C90538	hypothetical prote	1098	4	13.8	58	2	H86755	prophage p12 prote
1026	5	17.2	3655	2	T38084	TRAP-like protein	1099	4	13.8	59	1	T1EBED	venom basic protei
1027	5	17.2	3898	1	A44217	genome polyprotein	1100	4	13.8	59	2	A59399	long epsilon-dendr
1028	5	17.2	4549	2	T20771	hypothetical prote	1101	4	13.8	60	2	S36770	homeotic protein e
1029	5	17.2	4667	2	T20774	hypothetical prote	1102	4	13.8	60	2	S13129	probable homeotic
1030	5	17.2	4725	1	A44357	dynein heavy chain	1103	4	13.8	60	2	PC1228	homeotic protein C
1031	5	17.2	4919	2	T31105	hypothetical prote	1104	4	13.8	60	2	AC2576	hypothetical prote
1032	5	17.2	5147	1	LJFTFM	cadherin-related t	1105	4	13.8	61	4	S53070	hypothetical prote
1033	5	17.2	5825	2	T12117	polyprotein - fava	1106	4	13.8	62	2	G97332	hypothetical prote
1034	5	17.2	10797	2	T30192	probable peptide s	1107	4	13.8	63	2	S11883	copB protein - Esc
1035	5	17.2	13055	2	T16580	hypothetical prote	1108	4	13.8	63	2	T31193	hypothetical prote
1036	5	17.2	15281	2	S41309	cyclosporin synthe	1109	4	13.8	63	2	E82775	hypothetical prote
1037	4	13.8	12	5	A55837	5-aminoimidazole r	1110	4	13.8	63	2	D98278	hypothetical prote
1038	4	13.8	15	2	S08282	cytochrome P450K-2	1111	4	13.8	64	1	S03778	uviB protein - Clo
1039	4	13.8	16	2	PC4416	Erbb kinase activa	1112	4	13.8	64	2	F70188	hypothetical prote
1040	4	13.8	16	2	S24667	protein-tyrosine k	1113	4	13.8	64	2	T39013	hypothetical prote
1041	4	13.8	17	2	S17274	ribosomal protein	1114	4	13.8	64	2	F96006	hypothetical prote
1042	4	13.8	20	2	A41717	P100 protein - rat	1115	4	13.8	65	2	A90884	hypothetical prote
1043	4	13.8	25	2	S71387	alpha-2-macroglobu	1116	4	13.8	65	2	B45692	orf2 within p gene
1044	4	13.8	30	2	S25666	phosphopyruvate hy	1117	4	13.8	65	2	A81082	hypothetical prote
1045	4	13.8	30	2	S01657	atrial natriuretic	1118	4	13.8	65	2	H64895	hypothetical prote
1046	4	13.8	30	2	A84412	hypothetical prote	1119	4	13.8	65	2	S08014	calcium-binding pr
1047	4	13.8	31	2	D82093	hypothetical prote	1120	4	13.8	65	2	A71150	hypothetical prote
1048	4	13.8	33	2	S70495	ferritin light cha	1121	4	13.8	65	2	E82547	hypothetical prote
1049	4	13.8	33	2	F87516	hypothetical prote	1122	4	13.8	65	2	F85734	hypothetical prote
1050	4	13.8	34	2	C43853	phosphopyruvate hy	1123	4	13.8	66	2	S26663	microtubule-associ
1051	4	13.8	34	2	C49195	corticostatic pept	1124	4	13.8	66	2	D87482	conserved hypothet

1125 4 13.8 66 2 T28193 hypothetical prote
1126 4 13.8 67 1 B48180 small acid-soluble
1127 4 13.8 67 2 S43066 hypothetical prote
1128 4 13.8 67 2 B69830 hypothetical prote
1129 4 13.8 67 2 B83792 small acid-soluble
1130 4 13.8 68 2 S14331 major proteoglycan
1131 4 13.8 68 2 T30399 hypothetical prote
1132 4 13.8 68 2 T17853 hypothetical prote
1133 4 13.8 68 2 T45155 hypothetical prote
1134 4 13.8 68 2 T00189 hypothetical prote
1135 4 13.8 69 1 A48180 small acid-soluble
1136 4 13.8 69 2 D70891 hypothetical prote
1137 4 13.8 69 2 B69742 hypothetical prote
1138 4 13.8 69 2 C22112 hypothetical prote
1139 4 13.8 69 2 A11984 molybdo-pterin bin
1140 4 13.8 69 2 AG3599 transcription regu
1141 4 13.8 70 1 C25234 small acid-soluble
1142 4 13.8 70 2 A13015 hypothetical prote
1143 4 13.8 71 1 D25234 small acid-soluble
1144 4 13.8 71 2 I51606 gene c-jun protein
1145 4 13.8 71 2 S21364 dnaK-type molecula
1146 4 13.8 71 2 S08019 lysis protein - ph
1147 4 13.8 71 2 T19056 hypothetical prote
1148 4 13.8 71 2 AE2139 hypothetical prote
1149 4 13.8 71 2 AG1385 hypothetical prote
1150 4 13.8 72 2 G72399 ribosomal protein
1151 4 13.8 72 2 G83106 hypothetical prote
1152 4 13.8 72 2 C75519 hypothetical prote
1153 4 13.8 72 2 T28329 ORF MSV168 hypothe
1154 4 13.8 72 2 AE1875 hypothetical prote
1155 4 13.8 73 2 H69460 conserved hypotet
1156 4 13.8 73 2 S28978 srb protein - phag
1157 4 13.8 73 2 H81817 hypothetical prote
1158 4 13.8 73 2 D70869 hypothetical prote
1159 4 13.8 73 2 B71297 probable carbon st
1160 4 13.8 74 2 T17513 hypothetical prote
1161 4 13.8 74 2 G75443 hypothetical prote
1162 4 13.8 74 2 C69802 hypothetical prote
1163 4 13.8 75 1 YVBPM5 lysis protein - ph
1164 4 13.8 75 2 E75258 hypothetical prote
1165 4 13.8 75 2 T01993 hypothetical prote
1166 4 13.8 75 2 S45069 recombinant antige
1167 4 13.8 75 2 E95878 hypothetical prote
1168 4 13.8 75 2 G95891 hypothetical prote
1169 4 13.8 75 2 AC3575 hypothetical prote
1170 4 13.8 76 1 JC4128 cAMP-dependent pro
1171 4 13.8 76 2 I68914 MHC protein - cott
1172 4 13.8 76 2 I68913 MHC protein - cott
1173 4 13.8 76 2 I68912 MHC protein - cott
1174 4 13.8 76 2 C91035 hypothetical prote
1175 4 13.8 76 2 T48333 hypothetical prote
1176 4 13.8 76 2 AG2808 hypothetical prote
1177 4 13.8 77 2 S12514 conotoxin Tx-KK1 p
1178 4 13.8 77 2 I51318 bHLH transcription
1179 4 13.8 77 2 F87395 conserved tyrosine-
1180 4 13.8 78 2 S77855 probable tyrosine-
1181 4 13.8 78 2 H90114 putative small nuc
1182 4 13.8 78 2 C96701 unknown protein, 9
1183 4 13.8 78 2 T17804 hypothetical prote
1184 4 13.8 78 2 S70359 gene 15 protein ho
1185 4 13.8 78 2 AH1969 hypothetical prote
1186 4 13.8 78 2 AD2707 hypothetical prote
1187 4 13.8 79 1 TIEPVK K+ channel blocker
1188 4 13.8 79 2 JT0667 hypothetical 9.3K
1189 4 13.8 79 2 C73347 hypothetical prote
1190 4 13.8 79 2 E69854 hypothetical prote
1191 4 13.8 79 2 T35520 hypothetical prote
1192 4 13.8 79 2 AB1982 hypothetical prote
1193 4 13.8 80 2 A45940 calmodulin - sea u
1194 4 13.8 80 2 S21770 saposin-C - bovine
1195 4 13.8 80 2 S48628 hypothetical prote
1196 4 13.8 80 2 S19984 hypothetical prote
1197 4 13.8 80 2 S19987 hypothetical prote

1198 4 13.8 80 2 E82530 hypothetical prote
1199 4 13.8 81 1 A55790 ferredoxin [4Fe-4S
1200 4 13.8 81 2 H90872 hypothetical prote

ALIGNMENTS

RESULT 1
A47327
selenoprotein P precursor [validated] - human
C:Species: Homo sapiens (man)
C:Date: 21-Sep-1993 #sequence_revision 01-Dec-1995 #text_change 15-Sep-2000
C:Accession: A47327; S42752
R: Hill, K.E.; Lloyd, R.S.; Burk, R.F.
Proc. Natl. Acad. Sci. U.S.A. 90, 537-541, 1993
A:Title: Conserved nucleotide sequences in the open reading frame and 3' untranslated
A:Reference number: A47327; MUID:93133823
A:Accession: A47327
A:Molecule type: mRNA
A:Residues: 1-381 <HIL>
A:Cross-references: GB:Z11793; NID:q36425; PIDN:CAA77836.1; PID:q2654365
A:Experimental source: heart and liver
A:Note: in Genbank entry HSSELP, release 117.0, PIDN:CAA77836.1, the selenocysteine
R:Akesson, B.; Bellew, T.; Burk, R.F.
Biochim. Biophys. Acta 1204, 243-249, 1994
A:Title: Purification of selenoprotein P from human plasma.
A:Reference number: S42752; MUID:94191007
A:Accession: S42752
A:Molecule type: protein
A:Residues: 20-27,'X',29-33 <AKE>
A:Note: mature forms of 55K and 61K were detected in plasma; the protein was shown to
C:Genetics:
A:Gene: GDB:SEPP1; SLP
A:Cross-references: GDB:L38278; OMIM:601484
A:Map position: 5q31-5q31
C:Function:
A:Description: may act as a free-radical scavenger
C:Superfamily: selenoprotein P
C:Keywords: extracellular protein; glycoprotein; heparin binding; liver; plasma; sele
F:1-19/Domain: signal sequence #status predicted <Sig>
F:20-381/Product: selenoprotein P #status experimental <MAT>
F:46,83,119,128,338/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:59,300,318,330,345,352,367,369,376,378/Modified site: selenocysteine #status predic

Query Match 72.4%; Score 21; DB 1; Length 381;
Best Local Similarity 100.0%; Pred. No. 8,7e-15;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRCINQLCKLPTDSELAPRS 21
Db 279 KRCINQLCKLPTDSELAPRS 299
|||||

RESULT 2
T10442
selenoprotein P precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T10442
R:Steinert, P.; Ahrens, M.; Gross, G.; Flohe, L.
Biofactors 6, 311-319, 1997
A:Title: cDNA and deduced polypeptide sequence of a mouse selenoprotein P.
A:Reference number: Z17017; MUID:97434516
A:Accession: T10442
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-380 <STE>
A:Cross-references: EMBL:X99807; NID:q1495998; PIDN:CAA68140.1; PID:q1495999
C:Genetics:
A:Gene: selp
C:Superfamily: selenoprotein P

C;Keywords: extracellular protein; selenocysteine

Query Match 31.0%; Score 9; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CINQLLCKL 11
|||||
Db 281 CINQLLCKL 289

RESULT 3

OMRTSP

N;Alternate names: selenoprotein P precursor [validated] - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 15-Sep-2000
C;Accession: A40380; B40380; S68322
R;Hill, K.E.; Lloyd, R.S.; Yang, J.G.; Read, R.; Burk, R.F.
J. Biol. Chem. 266, 10050-10053, 1991
A;Title: The cDNA for rat selenoprotein P contains 10 TGA codons in the open reading frame
A;Reference number: A40380; MUID:91244760
A;Molecule type: mRNA
A;Accession: A40380
A;Residues: 1-385 <HIL>
A;Cross-references: GB:M63574; NID:g206893; PIDN:AAA42129.1; PID:g206894
A;Accession: B40380
A;Molecule type: protein
A;Residues: 20-41;267-287;316-327 <H12>
R;Chittum, H.S.; Himeno, S.; Hill, K.E.; Burk, R.F.
Arch. Biochem. Biophys. 325, 124-128, 1996
A;Title: Multiple forms of selenoprotein P in rat plasma.
A;Reference number: S68322; MUID:96140605
A;Accession: S68322
A;Molecule type: protein
A;Residues: 20-27 <CHI>

C;Superfamily: selenoprotein P
C;Keywords: extracellular protein; glycoprotein; heparin binding; liver; plasma; selenocysteine
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-385/Product: selenoprotein P #status experimental
F;59,264,335,357,371,373,380,382/Modified site: selenocysteine #status predicted
F;83,174,188,370,375/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;282,323/Modified site: selenocysteine #status experimental

Query Match 31.0%; Score 9; DB 1; Length 385;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CINQLLCKL 11
|||||
Db 286 CINQLLCKL 294

RESULT 4

AF2287

group 3 sigma 37-type sigma factor [imported] - Anabaena sp. (strain PCC 7120)
C;Species: Anabaena sp.
A;Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C;Accession: AF2287
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AF2287
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-258 <KUR>
A;Cross-references: GB:BA000019; PIDN:BA075552.1; PID:g17132987; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:

A;Gene: all3853

C;Superfamily: transcription sigma factor G; transcription initiation factor sigma ka

Query Match 24.1%; Score 7; DB 2; Length 258;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSLA 18
|||||
Db 132 PTDSLA 138

RESULT 5

A25872

N;Alternate names: transcription regulator GRF10 - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 21-Jul-2000
C;Accession: A25872; S67648; S19636; S67412; S72100
R;Sengstag, C.; Hinnen, A.
Nucleic Acids Res. 15, 233-246, 1987
A;Title: The sequence of the Saccharomyces cerevisiae gene PHO2 codes for a regulator
A;Reference number: A25872; MUID:871146334
A;Accession: A25872
A;Molecule type: DNA
A;Residues: 1-559 <SEN>
A;Cross-references: GB:M24613; NID:g4143; PIDN:CAA28729.1; PID:g4144
R;Ballesta, J.P.G.; Remacha, M.; Soler-Mira, A.; Jimenez, A.; Garcia-Cantalejo, J.M.;
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67629
A;Accession: S67648
A;Molecule type: DNA
A;Residues: 1-559 <BAL>
A;Cross-references: EMBL:Z74154; NID:g1431149; PIDN:CAA98673.1; PID:e2532225; PID:g143

A;Experimental source: strain S288C
Nucleic Acids Res. 19, 3463, 1991
A;Title: Molecular analysis of a temperature sensitive allele of the PHO2 gene of Sac
A;Reference number: S19636; MUID:91288241
A;Accession: S19636
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 65-156, 'E', 158-310, 'N', 312-559 <MCC>
A;Cross-references: EMBL:X54293; NID:g4146; PIDN:CAA38192.1; PID:g4147
A;Experimental source: strain GG100-14D

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1990
R;Boskovic, J.; Saiz, J.E.; Soler-Mira, A.; Garcia-Cantalejo, J.; Revuelta, J.L.; Jim
submitted to the EMBL Data Library, February 1996
A;Reference number: S67406
A;Accession: S67412
A;Molecule type: DNA
A;Residues: 1-559 <BOS>

A;Cross-references: EMBL:X95644; NID:g1199535; PIDN:CAA64906.1; PID:e223187; PID:g119
R;Saiz, J.E.; Buitrage, M.J.; Garcia, R.; Revuelta, J.L.; del Rey, F.
Yeast 12, 1077-1084, 1996

A;Title: The sequence of a 20.3 kb DNA fragment from the left arm of Saccharomyces ce
A;Reference number: S72094; MUID:97051597
A;Accession: S72100
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-559 <SAI>
A;Cross-references: EMBL:X95644; NID:g1199535; PIDN:CAA64906.1; PID:e223187; PID:g119
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1996
C;Genetics:
A;Gene: SGD:GRF10; PHO2; BAS2
A;Cross-references: MIPS:YDL106c; SGD:S0002264
A;Map position: 4L
C;Function:

A;Description: required for expression of phosphate pathway and other genes; acts as
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;78-134/Domain: homeobox homology <HOX>

Query Match 24.1%; Score 7; DB 2; Length 559;
 Best Local Similarity 100.0%; Pred. No. 7.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 LPTDSEL 17
 |||||
 Db 533 LPTDSEL 539

RESULT 6

T47327
 hypothetical protein T12K4.140 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T47327
 R:Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Mewes, H.W.; Rudd, S.; Lemcke, K.; May
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: Z24460
 A:Accession: T47327
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1009 <MON>
 A:Cross-references: EMBL:AL138640
 A:Experimental source: cultivar Columbia; BAC clone T12K4
 C:Genetics:
 A:Map position: 3
 A:Introns: 130/2; 250/3; 567/3; 608/3; 638/3; 681/3; 782/2; 810/3
 A:Note: T12K4.140

Query Match 24.1%; Score 7; DB 2; Length 1009;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KLPTDSE 16
 |||||
 Db 82 KLPTDSE 88

RESULT 7

D70026
 hypothetical protein yuze - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 28-Jul-2000
 C:Accession: D70026
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertoni
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
 A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrati, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinols,
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:98044033
 A:Accession: D70026
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-129 <KUN>
 A:Cross-references: GB:Z99120; GB:AL009126; NID:g2635613; PIDN:CAB15169.1; PID:ell184258;
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: yuze
 C:Superfamily: Bacillus subtilis hypothetical protein yuze

Query Match 20.7%; Score 6; DB 2; Length 129;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 TDSELA 18
 |||||
 Db 8 TDSELA 13

RESULT 8

C24925
 lactose permease - Klebsiella pneumoniae (fragment)
 N:Alternate names: galactose permease
 C:Species: Klebsiella pneumoniae
 C:Date: 23-Aug-1987 #sequence_revision 23-Aug-1987 #text_change 20-Aug-1999
 C:Accession: C24925
 R:Buvinger, W.E.; Riley, M.
 J. Bacteriol. 163, 850-857, 1985
 A:Title: Nucleotide sequence of Klebsiella pneumoniae lac genes.
 A:Reference number: A91803; MUID:85289025
 A:Accession: C24925
 A:Molecule type: DNA
 A:Residues: 1-133 <BOV>
 A:Cross-references: GB:M11441; GB:M11416; NID:g149216; PIDN:AAA25083.1; PID:g551868
 C:Genetics:
 A:Gene: lacy
 C:Superfamily: lactose permease
 C:Keywords: transmembrane protein

Query Match 20.7%; Score 6; DB 2; Length 133;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 SELAPR 20
 |||||
 Db 6 SELAPR 11

RESULT 9

RAWAI
 ribulose-bisphosphate carboxylase (EC 4.1.1.39) small chain A1 precursor - Arabidopsi
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 30-Jun-1993
 C:Accession: S03720
 R:Krebers, E.; Seurinck, J.; Hardies, L.; Cashmore, A.R.; Timko, M.P.
 Plant Mol. Biol. 11, 745-759, 1988
 A:Title: Four genes in two diverged subfamilies encode the ribulose-1,5-bisphosphate
 A:Reference number: S03717
 A:Accession: S03720
 A:Molecule type: DNA
 A:Residues: 1-180 <KRE>
 A:Cross-references: EMBL:X13611
 C:Genetics:
 A:Gene: atslA; rbcS
 A:Introns: 57/3; 102/3
 C:Superfamily: ribulose-bisphosphate carboxylase small chain
 C:Keywords: Calvin cycle; carbon dioxide fixation; carbon-carbon lyase; carboxy-lyase
 F:1-55/Domain: transit peptide (chloroplast) #status predicted <INP>
 F:56-180/Product: ribulose-bisphosphate carboxylase small chain A1 #status predicted

Query Match 20.7%; Score 6; DB 1; Length 180;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 TDSELA 18
 |||||
 Db 77 TDSELA 82

RESULT 10

G96694

hypothetical protein F5A8.1 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
 C:Accession: G96694
 R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.N.F.; Hughes, B.; Huizar, L.
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: G96694
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-180 <STO>
 A:Cross-references: GB:AE005173; MID:g4204274; PIDN:AAD10655.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F5A8.1
 A:Map position: 1
 C:Superfamily: ribulose-bisphosphate carboxylase small chain

Query Match 20.7%; Score 6; DB 2; Length 180;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDSELA 18
 |||||
 Db 77 TDSELA 82

RESULT 11
 A71177
 probable thymidylate kinase - Pyrococcus horikoshii
 C:Species: Pyrococcus horikoshii
 C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
 C:Accession: A71177
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekir M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi DNA Res. 5, 55-76, 1998
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic A:Reference number: A71000; MUID:98344137
 A:Accession: A71177
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-205 <RAW>
 A:Cross-references: GB:AP000007; MID:g3236134; PIDN:BAA30808.1; PID:g3258125
 A:Experimental source: strain OT3
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
 C:Genetics:
 A:Gene: PH1695
 C:Superfamily: dTMP kinase

Query Match 20.7%; Score 6; DB 2; Length 205;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSEL 17
 |||||
 Db 40 PTDSEL 45

RESULT 12

T36656
 probable carbonic anhydrase - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 19-May-2000

C:Accession: T36656
 R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A submitted to the EMBL Data Library, June 1999
 A:Reference number: Z21610
 A:Accession: T36656
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-207 <OLI>
 A:Cross-references: EMBL:AL078610; PIDN:CAB44424.1; GSPDB:GN00070; SCOEEDB:SCH35.03
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEEDB:SCH35.03
 C:Superfamily: Escherichia coli carbonate dehydratase

Query Match 20.7%; Score 6; DB 2; Length 207;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SELAPR 20
 |||||
 Db 39 SELAPR 44

RESULT 13

AI0870
 hypothetical protein STV3191 [imported] - Salmonella enterica subsp. enterica serovar C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
 C:Accession: AI0870
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se A:Reference number: AB0502; PMID:11677608
 A:Accession: AI0870
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-217 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD02865.1; PID:g16504118; GSPDB:GN00176
 C:Genetics:
 A:Gene: STV3191

Query Match 20.7%; Score 6; DB 2; Length 217;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 LPTDSE 16
 |||||
 Db 195 LPTDSE 200

RESULT 14

T20219
 hypothetical protein C5AD10.10 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T20219; T21470
 R:Dobson, R.
 submitted to the EMBL Data Library, June 1996
 A:Reference number: Z19238
 A:Accession: T20219
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-228 <WIL>
 A:Cross-references: EMBL:Z75531; PIDN:CAA99808.1; GSPDB:GN00023; CESP:C54D10.10
 A:Experimental source: clone C54D10
 R:Kershaw, J.
 submitted to the EMBL Data Library, June 1996

A:Reference number: Z19427
A:Accession: T21470
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-228 <W12>
A:Cross-references: EMBL:Z75539; PIDN:CAA99845.1; GSPDB:GN00023; CESP:C54D10.10
A:Experimental source: clone F28C1
C:Genetics:
A:Gene: CESP:C54D10.10
A:Map position: 5
A:Introns: 11/1; 28/1; 59/1; 130/1

Query Match 20.7%; Score 6; DB 2; Length 228;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 CKLPTD 14
|||||
DB 133 CKLPTD 138

RESULT 15

E90674
hypothetical protein ECS0365 [imported] - Escherichia coli (strain O157:H7, substrain R1
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: E90674
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kunara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: E90674
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-236 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA833788.1; PID:gl3359822; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECS0365

Query Match 20.7%; Score 6; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CINQLL 8
|||||
DB 100 CINQLL 105

RESULT 16

H85524
hypothetical protein Z0405 [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: H85524
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: H85524
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-236 <STO>
A:Cross-references: GB:AB005174; NID:gl2513133; PIDN:AAG54660.1; GSPDB:GN00145; UWGP:Z04
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z0405

Query Match 20.7%; Score 6; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CINQLL 8
|||||
DB 100 CINQLL 105

RESULT 17

B95267
probable ABC transporter ATP-binding protein SMA0083 [imported] - Sinorhizobium meli
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: B95267
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; B
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meli
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: B95267
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-259 <KUR>
A:Cross-references: GB:AE006469; PIDN:AAK64700.1; PID:gl4523100; GSPDB:GN00165
A:Experimental source: strain 1021, megaplasmid pSymA
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMA0083
A:Genome: plasmid

Query Match 20.7%; Score 6; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RCINQL 7
|||||
DB 64 RCINQL 69

RESULT 18

A82104
conserved hypothetical protein VC2229 [imported] - Vibrio cholerae (strain N16961 ser
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: A82104
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833
A:Accession: A82104
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-281 <HEI>
A:Cross-references: GB:AE004294; GB:AE003852; NID:g9656774; PIDN:AAF95373.1; GSPDB:GN
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC2229
A:Map position: 1
C:Superfamily: hypothetical protein HI1037

Query Match 20.7%; Score 6; DB 2; Length 281;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 TDSELA 18
 |||||
 Db 128 TDSELA 133

RESULT 19
 D70955
 hypothetical protein Rv3603c - Mycobacterium tuberculosis (strain H37RV)
 C;Species: Mycobacterium tuberculosis
 C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
 C;Accession: D70955
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
 A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A;Reference number: A70500; MUID:98295987
 A;Accession: D70955
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-303 <COL>
 A;Cross-references: GB:Z95557; GB:AL123456; NID:g3242276; PIDN:CAB08941.1; PID:e316966;
 A;Experimental source: strain H37RV
 C;Genetics:
 A;Gene: RV3603c

Query Match 20.7%; Score 6; DB 2; Length 303;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 TDSELA 18
 |||||
 Db 78 TDSELA 83

RESULT 20
 D64758
 probable transcription regulator yabB - Escherichia coli
 C;Species: Escherichia coli
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
 C;Accession: D64758
 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.; Science 277, 1453-1462, 1997
 A;Title: The complete genome sequence of Escherichia coli K-12.
 A;Reference number: A64720; MUID:97426617
 A;Accession: D64758
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-310 <BLAT>
 A;Cross-references: GB:AE000138; GB:U00096; NID:g1786501; PIDN:AACT3419.1; PID:g1786508;
 A;Experimental source: strain K-12, substrain MG1655
 C;Genetics:
 A;Gene: yabB
 C;Superfamily: probable transcription regulator ybbS
 C;Keywords: DNA binding; transcription regulation
 F;22-41/Region: helix-turn-helix motif

Query Match 20.7%; Score 6; DB 1; Length 310;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CINOLL 8
 |||||
 Db 100 CINOLL 105

RESULT 21
 AB0317
 lipolic acid synthetase [imported] - Yersinia pestis (strain CO92)
 C;Species: Yersinia pestis
 C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
 C;Accession: AB0317
 R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tithball, R.W.; Holden, M.T.G.; Prentice, M.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Li, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001
 A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A;Reference number: AB0001; MUID:21470413; PMID:11586360
 A;Accession: AB0317
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-321 <KUR>
 A;Cross-references: GB:AL590842; PIDN:CAC92841.1; PID:gl5980585; GSPDB:GN00175
 C;Genetics:
 A;Gene: lipA
 C;Superfamily: lipolic acid synthase

Query Match 20.7%; Score 6; DB 2; Length 321;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KLPTDS 15
 |||||
 Db 44 KLPTDS 49

RESULT 22
 F70666
 probable alcohol dehydrogenase - Mycobacterium tuberculosis (strain H37RV)
 C;Species: Mycobacterium tuberculosis
 C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C;Accession: F70666
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
 A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A;Reference number: A70500; MUID:98295987
 A;Accession: F70666
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-346 <COL>
 A;Cross-references: GB:Z83859; GB:AL123456; NID:g3261678; PIDN:CAB06125.1; PID:g17811
 A;Experimental source: strain H37RV
 C;Genetics:
 A;Gene: adhA
 C;Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology
 F;36-336/Domain: long-chain alcohol dehydrogenase homology <LADH>

Query Match 20.7%; Score 6; DB 2; Length 346;
 Best Local Similarity 100.0%; Pred. No. 58;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 DSELAP 19
 |||||
 Db 157 DSELAP 162

RESULT 23
 AF2505
 hypothetical protein all7222 [imported] - Anabaena sp. (strain PCC 7120) plasmid pCC7
 C;Species: Anabaena sp.
 A;Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
 C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
 C;Accession: AF2505

R.Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi Nakazaki, N.; Shimpō, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AF2505
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-348 <KUR>
 A:Cross-references: GB:BA000020; PIDN:BA078306.1; PID:g17135760; GSPDB:GN00180
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: all7222
 A:Genome: plasmid

Query Match 20.7%; Score 6; DB 2; Length 348;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 CINQLL 8
 Db 229 CINQLL 234
 RESULT 24
 T16350
 hypothetical protein F42G9.8 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jan-2000
 A:Accession: T16350
 R:Taich, A.
 submitted to the EMBL Data Library, March 1996
 A:Description: The sequence of C. elegans cosmid F42G9.
 A:Reference number: Z18498
 A:Accession: T16350
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-359 <TAI>
 A:Cross-references: EMBL:U00051; NID:g1216305; PID:g1216308; PIDN:AAA91354.1; CESP:F42G9
 A:Experimental source: strain Bristol N2
 C:Genetics:
 A:Gene: CESP:F42G9.8
 A:Introns: 28/2; 50/2; 113/2; 155/3; 220/3; 290/2
 C:Superfamily: Caenorhabditis elegans hypothetical protein F42G9.8

Query Match 20.7%; Score 6; DB 2; Length 359;
 Best Local Similarity 100.0%; Pred. No. 60;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 15 SELAPR 20
 Db 161 SELAPR 166
 RESULT 25
 C75597
 thymidine diphosphoglucose 4,6-dehydratase - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 A:Accession: C75597
 R:White, O.; Elsen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896
 A:Accession: C75597
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-364 <WHI>
 A:Cross-references: GB:AE001862; GB:AE001825; NID:g6460468; PIDN:AAF12268.1; PID:g6460506

A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DRA0041
 A:Map position: 2
 C:Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homolo

Query Match 20.7%; Score 6; DB 2; Length 364;
 Best Local Similarity 100.0%; Pred. No. 61;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 16 ELAPRS 21
 Db 155 ELAPRS 160

RESULT 26
 I54525
 leukemia-related protein - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
 A:Accession: I54525
 R:Jackson, S.H.; Malech, H.L.; Kozak, C.A.; Lomax, K.J.; Gallin, J.I.; Holland, S.M.
 Immunogenetics 39, 272-275, 1994
 A:Title: Cloning and functional expression of the mouse homologue of p47phox.
 A:Reference number: I54525; MUID:94164697
 A:Accession: I54525
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-388 <RES>
 A:Cross-references: GB:L11455; NID:g309422; PIDN:AAA50469.1; PID:g557868
 C:Superfamily: neutrophil cytosol factor 1; SH3 homology
 F.161-208/Domain: SH3 homology <SH31>

Query Match 20.7%; Score 6; DB 2; Length 388;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 10 KLPTDS 15
 Db 124 KLPTDS 129

RESULT 27
 S36186
 mitogen-activated protein kinase kinase (EC 2.7.1.-) 1 - African clawed frog
 N:Alternate names: ERK kinase 1; extracellular signal-regulated kinase activator kinase
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jun-2000
 A:Accession: S36186; S37632; S23389
 R:Kosako, H.; Nishida, E.; Gotoh, Y.
 EMBO J. 12, 787-794, 1993
 A:Title: cDNA cloning of MAP kinase kinase reveals kinase cascade pathways in yeasts
 A:Reference number: S36186; MUID:93178455
 A:Accession: S36186
 A:Molecule type: mRNA
 A:Residues: 1-395 <KOS>
 A:Cross-references: EMBL:D13700; NID:g222964; PIDN:BA002860.1; PID:g222965
 A:Accession: S37632
 A:Molecule type: protein
 A:Residues: 2-20; 37-47; 49-57; 71-84; 169-177; 186-214; 282-287; 317-324; 327-341; 356-363 <K
 R:Kosako, H.; Gotoh, Y.; Matsuda, S.; Ishikawa, M.; Nishida, E.
 EMBO J. 11, 2903-2908, 1992

EMBO J. 11, 2903-2908, 1992
 A:Title: Xenopus MAP kinase activator is a serine/threonine/tyrosine kinase activated
 A:Reference number: S23389; MUID:92347324
 A:Accession: S23389
 A:Molecule type: protein
 A:Residues: 186-203, 'X', 205-206, 'X', 208-214 <KO3>
 C:Complex: monomer
 C:Function:
 A:Description: catalyzes the formation of specific peptidyl-threonine-phosphate and p
 A:Pathway: growth factor-stimulated cascade in which Raf-1, MAPKK, and MAPK are sequ

C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine/tyrosine-specific
F:2-395/Product: mitogen-activated protein kinase kinase #status experimental <MAT>
F:66-363/Domain: protein kinase homology <KIN>
F:74-82/Region: protein kinase ATP-binding motif
F:97,114,190,192/Active site: Lys, Glu, Asp, Lys #status predicted
F:388/Binding site: phosphate (Thr) (covalent) (by MAP kinase) #status predicted

Query Match 20.7%; Score 6; DB 1; Length 395;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SELAPR 20
Db 286 SELAPR 291

RESULT 28

JT0487
lactose permease - Klebsiella pneumoniae (strain M5a1)
N:Alternate names: lactose transport protein
C:Species: Klebsiella pneumoniae
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Feb-1997
R:McMorrow, I.; Chin, D.T.; Fiebig, K.; Pierce, J.L.; Wilson, D.M.; Reeve, E.C.R.; Wilso
Biochim. Biophys. Acta 945, 315-323, 1988
A:Title: The lactose carrier of Klebsiella pneumoniae M5a1; the physiology of transport
A:Reference number: JT0487; MUID:89050993
A:Accession: JT0487
A:Molecule type: DNA
A:Residues: 1-416 <MCM>
C:Genetics:
A:Gene: lacY
C:Superfamily: lactose permease
C:Keywords: sugar transport; transmembrane protein

Query Match 20.7%; Score 6; DB 2; Length 416;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SELAPR 20
Db 4 SELAPR 9

RESULT 29

G85059
probable sugar transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001
C:Accession: G85059
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sprin
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488
A:Accession: G85059
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-461 <SPT>
A:Cross-references: GB:NC_001268; NID:g7267233; PTID:G850840.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4g04750
A:Map position: 4
C:Superfamily: glucose transport protein

Query Match 20.7%; Score 6; DB 2; Length 461;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SELAPR 20

Db 131 SELAPR 136

RESULT 30

JK0201
alpha-amylase (EC 3.2.1.1) - Aspergillus oryzae
N:Alternate names: glycogenase; Taka-amylase A
C:Species: Aspergillus oryzae
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998
C:Accession: JK0201
R:Toda, H.; Kondo, K.; Narita, K.
Proc. Jpn. Acad. 58B, 208-212, 1982
A:Title: The complete amino acid sequence of Taka-amylase A.
A:Reference number: JK0201
A:Accession: JK0201
A:Molecule type: protein
A:Residues: 1-478 <TOD>
C:Comment: One atom of calcium per molecule is essential for the activity.
C:Comment: This enzyme is a glycoprotein.
C:Comment: See also PIR:JT0466 and PIR:JS0240.
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:173-300/Domain: alpha-amylase core homology <AMY>
F:197/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:210,230,297/Active site: His, Glu, Asp #status experimental

Query Match 20.7%; Score 6; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSSEL 17

Db 353 PTDSSEL 358

RESULT 31

A48305
alpha-amylase (EC 3.2.1.1) A precursor - Aspergillus awamori
C:Species: Aspergillus awamori
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997
C:Accession: A48305
R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J
Curr. Genet. 17, 203-212, 1990
A:Title: Cloning, characterization, and expression of two alpha-amylase genes from As
A:Reference number: A48305; MUID:90254827
A:Accession: A48305
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-498 <KOR>
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 20.7%; Score 6; DB 2; Length 498;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSSEL 17

Db 374 PTDSSEL 379

RESULT 32

ALAS1

alpha-amylase (EC 3.2.1.1) 1 precursor [validated] - Aspergillus oryzae
N:Alternate names: alpha-amylase isozyme II; glycogenase; Taka-amylase A
C:Species: Aspergillus oryzae
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 15-Sep-2000
C:Accession: S04548; A33214; J50240; A91930; A93767; A10627
R:Wirsal, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
Mol. Microbiol. 3, 3-14, 1989

A:Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exon
A:Reference number: S04548; MUID:89237897

A:Accession: S04548

A:Molecule type: DNA

A:Residues: 1-499 <WIR>

A:Cross-references: EMBL:X12725; NID:g2430; PIDN:CAA31218.1; PID:g295921

A:Genetics: AMY1

A:Accession: A33214

A:Molecule type: mRNA

A:Residues: 1-499 <W12>

A:Cross-references: GB:X12725; NID:g2430; PIDN:CAA31218.1; PID:g295921

R:Genes, M.J.; Dove, M.J.; Seligy, V.L.

Gene 79, 107-117, 1989

A:Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each containing
A:Reference number: J50240; MUID:89378767

A:Accession: J50240

A:Molecule type: DNA

A:Residues: 1-499 <GEN>

A:Genetics: AMY2

A:Note: the authors refer to this as isozyme II

R:Isamura, S.; Ikenaka, T.

J. Biochem. 74, 1-10, 1973

A:Reference number: A91930; MUID:74001521

A:Accession: A91930

A:Molecule type: protein

A:Residues: 206-225 <ISE>

R:Narita, K.

Proc. Jpn. Acad. 51, 285-290, 1975

A:Reference number: A93767

A:Accession: A93767

A:Molecule type: protein

A:Residues: 434-443, 446-447, 'O', 449-458, 'GTTV', 459-464, 467-468, 'B', 470, 'B', 472-499 <NAR>
R:Matsuura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.

J. Biochem. 95, 697-702, 1984

A:Title: Structure and possible catalytic residues of Taka-amylase A.

A:Reference number: A37454; MUID:84212370

A:Contents: annotation; X-ray crystallography, 3.0 angstroms
R:Swift, H.J.; Brady, L.; Derewenda, Z.S.; Dodson, E.J.; Turkenburg, J.P.; Wilkinson, A.

submitted to the Brookhaven Protein Data Bank, August 1992

A:Reference number: A51548; PDB:6TAA

A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 22-497

C:Comment: One atom of calcium per molecule is essential for activity.

C:Genetics: <AMY1>

A:Gene: amy1

A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3

C:Genetics: <AMY2>

A:Gene: amy2; AmyII

A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation

C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology

C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; polys
F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-499/Product: alpha-amylase 1 #status experimental

F:194-321/Domain: alpha-amylase core homology <MAT>

F:51-59,171-185,261-304,461-496/Disulfide bonds: #status experimental

F:142,163,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted

F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:227,251,318/Active site: Asp, Glu, Asp #status predicted

Query Match 20.7%; Score 6; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 PTDSEL 17
|||||
Db 374 PTDSEL 379

RESULT 33

ALAS3

alpha-amylase (EC 3.2.1.1) 3 precursor - Aspergillus oryzae

N:Alternate names: alpha-amylase isozyme I; glycogenase; Taka-amylase A

C:Species: Aspergillus oryzae

C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Jun-1999

C:Accession: S04549; A33215; A44713

R:Wirsal, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.

Mol. Microbiol. 3, 3-14, 1989

A:Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exo

A:Reference number: S04548; MUID:89237897

A:Accession: S04549

A:Molecule type: DNA

A:Residues: 1-499 <WIR>

A:Cross-references: EMBL:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922

A:Accession: A33215

A>Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-499 <W12>

A:Cross-references: GB:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922

R:Genes, M.J.; Dove, M.J.; Seligy, V.L.

Gene 79, 107-117, 1989

A:Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each contain

A:Reference number: J50240; MUID:89378767

A:Accession: A44713

A:Molecule type: DNA

A:Residues: 1-499 <GEN>

A:Note: the authors refer to this as isozyme I

R:Matsuura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.

J. Biochem. 95, 697-702, 1984

A:Title: Structure and possible catalytic residues of Taka-amylase A.

A:Reference number: A37454; MUID:84212370

A:Contents: annotation; X-ray crystallography, 3.0 angstroms

C:Comment: One atom of calcium per molecule is essential for activity.

C:Genetics:

A:Gene: amy3; AmyI

A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation

C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology

C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; me

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-499/Product: alpha-amylase 3 #status experimental

F:194-321/Domain: alpha-amylase core homology <MAT>

F:51-59,171-185,261-304,461-496/Disulfide bonds: #status experimental

F:142,163,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted

F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:231,251,318/Active site: His, Glu, Asp #status experimental

Query Match 20.7%; Score 6; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 PTDSEL 17
|||||
Db 374 PTDSEL 379

RESULT 34

B48305

alpha-amylase (EC 3.2.1.1) B precursor - Aspergillus awamori

C:Species: Aspergillus awamori

C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997

C:Accession: B48305

R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J

Curr. Genet. 17, 203-212, 1990

A:Title: Cloning, characterization, and expression of two alpha-amylase genes from *Aspergillus*
A:Reference number: A48305; MUID:90254827

A:Accession: B48305
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-499 <KOR>

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
A:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 20.7%; Score 6; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSEL 17

Db 374 PTDSEL 379

RESULT 35

JS0663

alpha-amylase (EC 3.2.1.1) precursor - *Aspergillus* sp.

C:Species: *Aspergillus* sp.

C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 17-Mar-1999

C:Accession: JS0663

R:Shibuya, I.; Tamura, G.; Ishikawa, T.; Hara, S.

Biosci. Biotechnol. Biochem. 56, 174-179, 1992

A:Title: Cloning of the alpha-amylase cDNA of *Aspergillus shiroyamii* and its expression
A:Reference number: JS0663; MUID:92323146

A:Accession: JS0663

A:Molecule type: mRNA

A:Residues: 1-499 <SHI>

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation

C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology

C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-499/Product: alpha-amylase #status predicted <ALD>

F:194-321/Domain: alpha-amylase core homology <AMY>

F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 20.7%; Score 6; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSEL 17

Db 374 PTDSEL 379

RESULT 36

JT0466

alpha-amylase (EC 3.2.1.1) precursor - *Aspergillus oryzae*

N:Alternate names: glycogenase; taka-amylase A

C:Species: *Aspergillus oryzae*

C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998

C:Accession: JT0466

R:Tada, S.; Iimura, Y.; Gomi, K.; Takahashi, K.; Hara, S.; Yoshizawa, K.

Agric. Biol. Chem. 53, 593-599, 1989

A:Title: Cloning and nucleotide sequence of the genomic Taka-amylase A gene of *Aspergillus*

A:Reference number: JT0466

A:Accession: JT0466

A:Molecule type: DNA

A:Residues: 1-499 <TAD>

C:Comment: See also PIR:JK0201 and PIR:JS0240.

C:Function: One atom of calcium per molecule is essential for activity.

C:Genetics:

A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
A:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-499/Product: alpha-amylase #status predicted <NAV>

F:194-321/Domain: alpha-amylase core homology <AMY>

F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:231-251,318/Active site: His, Glu, Asp #status predicted

Query Match 20.7%; Score 6; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSEL 17

Db 374 PTDSEL 379

RESULT 37

JN0588

alpha-amylase (EC 3.2.1.1) precursor - *Aspergillus oryzae*

N:Alternate names: Taka-amylase A

C:Species: *Aspergillus oryzae*

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 14-Nov-1997

C:Accession: JN0588

R:Tsukagoshi, N.; Furukawa, M.; Nagaba, H.; Kirita, N.; Tsuboi, A.; Udaka, S.

Gene 84, 319-327, 1989

A:Title: Isolation of a cDNA encoding *Aspergillus oryzae* Taka-amylase A: evidence for
A:Reference number: JN0588; MUID:90128276

A:Accession: JN0588

A:Molecule type: mRNA

A:Residues: 1-499 <TSU>

C:Comment: The alpha amylases are encoded by multigene family.

C:Genetics:

A:Genes: Taa-GI

A:Introns: 56/1; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation

C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology

C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-499/Product: alpha-amylase #status predicted <NAV>

F:194-321/Domain: alpha-amylase core homology <AMY>

F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 20.7%; Score 6; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSEL 17

Db 374 PTDSEL 379

RESULT 38

AD1615

glycine betaine/carnitine/choline ABC transporter (membrane protein) homolog lin1461

C:Species: *Listeria innocua*

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C:Accession: AD1615

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl,

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;

Ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A:Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AD1615
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-504 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC96692.1; PID:g16413934; GSPDB:GN00178
A:Experimental source: strain Clp11262
C:Genetics:
A:Gene: lin1461

Query Match 20.7%; Score 6; DB 2; Length 504;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDSELA 18
|||||
Db 420 TDSELA 425

RESULT 39
AF1252
glycine betaine/carnitine/choline ABC transporter (membrane protein) homolog lmo1422 [im
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AF1252
R:Glaser, P.; Prangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maltournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant,
A:Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AF1252
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-504 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAC99500.1; PID:g16410851; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo1422

Query Match 20.7%; Score 6; DB 2; Length 504;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDSELA 18
|||||
Db 420 TDSELA 425

RESULT 40
AB0528
deoxyguanosinetriphosphate triphosphohydrolase [imported] - *Salmonella enterica* subsp. e
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
A:Note: this species has also been called *Salmonella typhi*
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: AB0528
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar
.; S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov
A:Reference number: AB0502; PMID:11677608
A:Accession: AB0528
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-505 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD01362.1; PID:g16501489; GSPDB:GN00176
C:Genetics:

A:Gene: STY0230

Query Match 20.7%; Score 6; DB 2; Length 505;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KLPTDS 15
|||||
Db 461 KLPTDS 466

RESULT 41
T22803
hypothetical protein F56H6.10 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22803
R:Kershaw, J.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19617
A:Accession: T22803
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-609 <WIL>
A:Cross-references: EMBL:Z81553; PIDN:CAB04499.1; GSPDB:GN00019; CESP:F56H6.10
C:Genetics:
A:Gene: CESP:F56H6.10
A:Map position: 1
A:Introns: 12/3; 48/2; 155/2; 224/1; 251/3; 434/2; 474/2; 520/2; 544/1; 555/2; 579/1

Query Match 20.7%; Score 6; DB 2; Length 609;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CINQLL 8
|||||
Db 122 CINQLL 127

RESULT 42
T22716
hypothetical protein F55C5.8 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22716
R:Harris, B.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z19604
A:Accession: T22716
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-622 <WIL>
A:Cross-references: EMBL:Z78198; PIDN:CAB01573.1; GSPDB:GN00023; CESP:F55C5.8
A:Experimental source: clone F55C5
C:Genetics:
A:Gene: CESP:F55C5.8
A:Map position: 5
A:Introns: 24/3; 44/2; 147/3; 251/3; 561/3

Query Match 20.7%; Score 6; DB 2; Length 622;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDSELA 18
|||||
Db 499 TDSELA 504

RESULT 43

T24486
Hypothetical protein T05A1.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24486
R:Lloyd, C.
submitted to the EMBL Data Library, December 1995
A:Reference number: Z19897
A:Accession: T24486
A:Status: preliminary; translated from GR/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-683 <WIL>
A:Cross-references: EMBL:Z68219; PIDN:CAA92480.1; GSPDB:GN00022; CESP:T05A1.3
A:Experimental source: clone T05A1
C:Genetics:
A:Map position: 4
A:Introns: 54/3; 115/3; 165/2; 207/2; 293/2; 336/3; 375/3; 411/3; 448/3; 473/3; 531/3; 6
Query Match 20.7%; Score 6; DB 2; Length 683;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 NQLLCK 10
DB 367 NQLLCK 372
RESULT 44
T09938
Hypothetical protein T16L4.260 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 15-Oct-1999
C:Accession: T09938
R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z16897
A:Accession: T09938
A:Molecule type: DNA
A:Residues: 1-776 <BEV>
A:Cross-references: EMBL:AL079344; GSPDB:GN00062; ATSP:T16L4.260
A:Experimental source: cultivar Columbia; BAC clone T16L4
C:Genetics:
A:Gene: ATSP:T16L4.260
A:Map position: 4
A:Introns: 308/3; 419/1; 462/3; 482/1; 574/3; 627/1; 724/3; 744/3
Query Match 20.7%; Score 6; DB 2; Length 776;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 LPTDSE 16
DB 598 LPTDSE 603
RESULT 45
A44048
genome polyprotein - Vilyuisk virus (strain V-1) (fragment)
C:Species: Vilyuisk virus
C:Date: 11-Feb-1993 #sequence_revision 11-Feb-1993 #text_change 07-Feb-1997
C:Accession: A44048
R:Pritchard, A.E.; Strom, T.; Lipton, H.L.
Virology 191, 469-472, 1992
A:Title: Nucleotide sequence identifies Vilyuisk virus as a divergent Theiler's virus.
A:Reference number: A44048; MUID:93053144
A:Accession: A44048
A:Molecule type: genomic RNA
A:Residues: 1-929 <PRI>
A:Cross-references: GB:M94868
C:Superfamily: foot-and-mouth disease virus genome polyprotein

C:Keywords: polyprotein
Query Match 20.7%; Score 6; DB 2; Length 929;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 LCKLPT 13
DB 461 LCKLPT 466
RESULT 46
S27756
receptor-like protein kinase 5 (EC 2.7.1.1-) precursor - Arabidopsis thaliana
N:Alternate names: protein F2009.180
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 11-Jun-1999
C:Accession: S27756; T04620
R:Walker, J.C.
submitted to the EMBL Data Library, February 1992
A:Description: Receptor-like protein kinase genes of Arabidopsis thaliana.
A:Reference number: S27754
A:Accession: S27756
A:Molecule type: mRNA
A:Residues: 1-999 <WAL>
A:Cross-references: EMBL:M84660; NID:g166849; PIDN:AAA32859.1; PID:g166850
A:Experimental source: strain Columbia
R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Hoheisel, J.; Mewes, H.W.; Mayer, K.
submitted to the Protein Sequence Database, October 1998
A:Reference number: Z15380
A:Accession: T04620
A:Molecule type: DNA
A:Residues: 1-999 <BEV>
A:Cross-references: EMBL:AL021749
A:Experimental source: cultivar Columbia; BAC clone F2009
C:Genetics:
A:Gene: RLK5
A:Map position: 4
A:Introns: 868/1
A:Note: F2009.180
C:Superfamily: protein kinase Xa21; leucine-rich alpha-2-glycoprotein repeat homology
C:Keywords: AIP; autophosphorylation; glycoprotein; kinase #status predicted <SIG>
F:1-20/Domain: signal sequence #status predicted <SIG>
F:66-89/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F:90-114/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F:115-139/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F:140-163/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F:184-187/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F:188-212/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F:213-236/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F:237-260/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
F:261-284/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
F:285-307/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>
F:308-331/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>
F:332-355/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>
F:356-379/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>
F:380-403/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>
F:404-427/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>
F:428-451/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR16>
F:452-475/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR17>
F:476-499/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR18>
F:500-523/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR19>
F:524-547/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR20>
F:548-570/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR21>
F:571-593/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR22>
F:681-971/Domain: protein kinase homology <KIN>
F:689-697/Region: protein kinase ATP-binding motif
F:98,102,150,185,210,269,282,452,576/Binding site: Lys, Glu, Asp, Lys #status predicted
F:711,737,819,821/Active site: Lys, Glu, Asp, Asp #status predicted
F:824,828/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 20.7%; Score 6; DB 1; Length 999;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 PTDSEL 17
 |||||
 Db 897 PTDSEL 902

RESULT 47

S28104
 probable DNA-directed RNA polymerase (EC 2.7.7.6) - gill mushroom (Agaricus bitorquis) F
 C:Species: Agaricus bitorquis
 C:Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 23-Apr-1999
 C:Accession: S28104
 R:Robison, M.M.; Royer, J.C.; Horgen, P.A.
 Curr. Genet. 19, 495-502, 1991
 A:Title: Homology between mitochondrial DNA of Agaricus bisporus and an internal portion
 A:Reference number: S28103; MUID:91347410
 A:Accession: S28104
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1102 <ROB>
 A:Cross-references: EMBL:Z11513
 C:Genetics:
 A:Genome: plasmid
 C:Keywords: mitochondrion; nucleotidyltransferase; transcription

Query Match 20.7%; Score 6; DB 2; Length 1102;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KLPTDS 15
 |||||
 Db 933 KLPTDS 938

RESULT 48

S03722
 DNA-directed DNA polymerase (EC 2.7.7.7) - Neurospora intermedia mitochondrion plasmid (C
 C:Species: mitochondrion Neurospora intermedia
 C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Jun-2000
 C:Accession: S03722
 R:Pande, S.; Lemire, E.G.; Nargang, F.E.
 Nucleic Acids Res. 17, 2023-2042, 1989
 A:Title: The mitochondrial plasmid from Neurospora intermedia strain Labelle-1b contains
 eins.
 A:Reference number: S03722; MUID:89183609
 A:Accession: S03722
 A:Molecule type: DNA
 A:Residues: 1-1151 <PAN>
 A:Cross-references: EMBL:X13912; NID:G2942; PIDN:CAB40819.1; PID:G4584824
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SSC3
 C:Keywords: DNA biosynthesis; mitochondrion; nucleotidyltransferase

Query Match 20.7%; Score 6; DB 2; Length 1151;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRCINQ 6
 |||||
 Db 374 KRCINQ 379

RESULT 49

S75705
 hypothetical protein sll0178 - Synechocystis sp. (strain PCC 6803)
 C:Species: Synechocystis sp.

A:Variety: PCC 6803
 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
 C:Accession: S75705
 R:Kaneke, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,
 O. K.; Okumura, S.; Shimpou, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
 DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
 S.
 A:Reference number: S74322; MUID:97061201
 A:Accession: S75705
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1319 <KAN>
 A:Cross-references: EMBL:D64003; GB:AB001339; NID:G1001200; PIDN:BAAL0440.1; PID:G100
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C:Superfamily: Synechocystis hypothetical protein sll0178

Query Match 20.7%; Score 6; DB 2; Length 1319;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 PTDSEL 17
 |||||
 Db 430 PTDSEL 435

RESULT 50

GNNYTN
 genome polypotein - murine poliovirus (strain DA)
 N:Contains: coat protein VP1; coat protein VP2; coat protein VP3; coat protein VP4; c
 -): protein p3-1b; RNA-directed RNA polymerase (EC 2.7.7.48)
 C:Species: murine poliovirus, Theiler's encephalomyelitis virus
 C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 16-Jul-1999
 C:Accession: A31228
 R:Ohara, Y.; Stein, S.; Fu, J.; Stillman, L.; Klamman, L.; Roos, R.P.
 Virology 164, 245-255, 1988
 A:Title: Molecular cloning and sequence determination of DA strain of Theiler's murin
 A:Reference number: A31228; MUID:88206072
 A:Accession: A31228
 A:Molecule type: genomic RNA
 A:Residues: 1-2301 <OHA>
 A:Cross-references: GB:M20301; NID:G335219; PIDN:AAA47928.1; PID:G335220
 C:Superfamily: foot-and-mouth disease virus genome polypotein
 C:Keywords: coat protein; core protein; genome-linked protein; hydrolase; nucleotidyl
 F:1-76/Domain: leader peptide #status predicted <LDP>
 F:77-147/Product: protein 1A (coat protein VP4) #status predicted <VP4>
 F:148-414/Product: protein 1B (coat protein VP2) #status predicted <VP2>
 F:415-646/Product: protein 1C (coat protein VP3) #status predicted <VP3>
 F:647-920/Product: protein 1D (coat protein VP1) #status predicted <VP1>
 F:921-1062/Product: protein 2A (core protein P2-3b) #status predicted <P2A>
 F:1063-1189/Product: protein 2B (core protein P2-5b) #status predicted <P2B>
 F:1190-1515/Product: protein 2C (core protein P2-X) #status predicted <P2C>
 F:1516-1603/Product: protein 3A (protein P3-1b) #status predicted <P3A>
 F:1604-1623/Product: protein 3B (genome-linked protein VPg) #status predicted <P3B>
 F:1624-1840/Product: protein 3C (probable proteinase) #status predicted <P3C>
 F:1841-2301/Product: protein 3D (probable RNA-directed RNA polymerase) #status predic

Query Match 20.7%; Score 6; DB 1; Length 2301;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 LCKLPT 13
 |||||
 Db 461 LCKLPT 466

Search completed: August 22, 2002, 11:18:19
 Job time: 86 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 11:16:53 ; Search time 21.52 Seconds
(without alignments)
32.916 Million cell updates/sec

Title: US-09-856-199-1

Perfect score: 29

Sequence: 1 KRCINQLCKLPTDSELAPRSCCHRH 29

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1200 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	20.7	359	3	US-09-150-133-11
2	6	20.7	359	3	US-09-150-141-11
3	6	20.7	359	4	US-09-374-493-11
4	6	20.7	359	4	US-09-374-824-11
5	6	20.7	359	4	US-09-374-492-11
6	6	20.7	478	1	US-08-720-899-7
7	6	20.7	478	1	US-08-459-610-7
8	6	20.7	478	2	US-08-343-804-7
9	6	20.7	478	2	US-08-339-715A-2
10	6	20.7	478	2	US-08-600-908A-10
11	6	20.7	478	3	US-08-683-838A-10
12	6	20.7	478	4	US-08-182-859-7
13	6	20.7	616	1	US-08-453-895A-115
14	6	20.7	616	1	US-08-268-161A-115
15	6	20.7	616	2	US-08-453-702A-115
16	6	20.7	616	4	US-09-099-639-115
17	6	20.7	616	5	PCT-US95-08071-115
18	6	20.7	999	2	US-08-473-553A-5
19	5	17.2	15	4	US-08-602-999A-395
20	5	17.2	19	2	US-08-768-147B-5
21	5	17.2	19	4	US-09-107-149-6
22	5	17.2	35	4	US-09-082-279B-835
23	5	17.2	35	4	US-09-315-304B-835
24	5	17.2	51	4	US-08-905-223-384
25	5	17.2	99	2	US-09-047-125-5
26	5	17.2	99	3	US-07-736-335E-5
27	5	17.2	117	2	US-08-505-617-6

101	5	17.2	425	4	US-09-020-683-5	Sequence 5, Appl	174	5	17.2	592	4	US-08-845-258-24	Sequence 24, Appl
102	5	17.2	427	4	US-09-364-230-32	Sequence 32, Appl	175	5	17.2	592	4	US-08-990-571-24	Sequence 24, Appl
103	5	17.2	435	1	US-08-259-148A-18	Sequence 18, Appl	176	5	17.2	592	4	US-08-723-142A-24	Sequence 24, Appl
104	5	17.2	435	1	US-08-484-054-18	Sequence 18, Appl	177	5	17.2	597	6	5171850-2	Patent No. 5171850
105	5	17.2	435	2	US-07-876-941A-18	Sequence 18, Appl	178	5	17.2	603	3	US-08-482-677-8	Sequence 8, Appl
106	5	17.2	443	3	US-09-120-365-69	Sequence 69, Appl	179	5	17.2	604	2	US-08-635-137-2	Sequence 2, Appl
107	5	17.2	443	4	US-09-155-855-1	Sequence 1, Appl	180	5	17.2	604	4	US-09-136-981-2	Sequence 2, Appl
108	5	17.2	443	4	US-09-155-855-2	Sequence 2, Appl	181	5	17.2	605	1	US-08-152-019A-26	Sequence 26, Appl
109	5	17.2	443	4	US-09-515-039-69	Sequence 69, Appl	182	5	17.2	605	3	US-08-482-677-4	Sequence 4, Appl
110	5	17.2	443	4	US-09-543-744-1	Sequence 1, Appl	183	5	17.2	605	3	US-08-650-599A-1	Sequence 1, Appl
111	5	17.2	443	4	US-09-543-744-2	Sequence 2, Appl	184	5	17.2	605	4	US-09-490-517-1	Sequence 1, Appl
112	5	17.2	449	3	US-08-462-351-2	Sequence 2, Appl	185	5	17.2	606	3	US-08-665-259-22	Sequence 22, Appl
113	5	17.2	449	6	5194425-1	Patent No. 5194425	186	5	17.2	606	3	US-08-762-500-22	Sequence 22, Appl
114	5	17.2	449	6	5468481-1	Patent No. 5468481	187	5	17.2	606	4	US-09-306-902A-9	Sequence 9, Appl
115	5	17.2	453	1	US-08-769-309A-14	Sequence 14, Appl	188	5	17.2	616	4	US-09-268-347-38	Sequence 38, Appl
116	5	17.2	453	3	US-08-994-570-14	Sequence 14, Appl	189	5	17.2	632	1	US-08-221-817-11	Sequence 11, Appl
117	5	17.2	457	3	US-08-278-635B-3	Sequence 3, Appl	190	5	17.2	632	1	US-08-454-439-11	Sequence 11, Appl
118	5	17.2	457	3	US-08-471-961-3	Sequence 3, Appl	191	5	17.2	632	5	PCT-US94-10487-11	Sequence 11, Appl
119	5	17.2	458	3	US-08-464-258B-3	Sequence 3, Appl	192	5	17.2	634	1	US-07-779-049-3	Sequence 3, Appl
120	5	17.2	460	2	US-08-933-821-17	Sequence 17, Appl	193	5	17.2	634	1	US-08-080-240-3	Sequence 3, Appl
121	5	17.2	460	3	US-08-934-494-6	Sequence 6, Appl	194	5	17.2	651	1	US-08-769-309A-17	Sequence 17, Appl
122	5	17.2	460	3	US-08-960-507-17	Sequence 17, Appl	195	5	17.2	651	3	US-08-994-570-17	Sequence 17, Appl
123	5	17.2	460	3	US-09-143-068-6	Sequence 6, Appl	196	5	17.2	659	1	US-08-240-049B-16	Sequence 16, Appl
124	5	17.2	460	4	US-09-143-707-6	Sequence 6, Appl	197	5	17.2	659	1	US-08-259-148A-20	Sequence 20, Appl
125	5	17.2	460	4	US-09-202-089-6	Sequence 6, Appl	198	5	17.2	659	2	US-08-484-054-20	Sequence 20, Appl
126	5	17.2	460	4	US-09-136-828-17	Sequence 17, Appl	199	5	17.2	659	2	US-07-876-941A-20	Sequence 20, Appl
127	5	17.2	467	1	US-08-151-574-32	Sequence 32, Appl	200	5	17.2	659	4	US-08-477-292-14	Sequence 14, Appl
128	5	17.2	467	1	US-08-146-424-20	Sequence 20, Appl	201	5	17.2	659	4	US-09-268-347-46	Sequence 46, Appl
129	5	17.2	467	1	US-08-693-709-2	Sequence 2, Appl	202	5	17.2	660	4	US-08-542-634-14	Sequence 14, Appl
130	5	17.2	467	1	US-07-923-724-8	Sequence 8, Appl	203	5	17.2	660	4	US-09-268-347-45	Sequence 45, Appl
131	5	17.2	467	4	US-08-609-426A-8	Sequence 8, Appl	204	5	17.2	660	5	PCT-US95-13703-14	Sequence 14, Appl
132	5	17.2	467	2	US-08-374-652C-2	Sequence 2, Appl	205	5	17.2	664	1	US-08-421-661-6	Sequence 6, Appl
133	5	17.2	467	2	US-08-419-448-32	Sequence 32, Appl	206	5	17.2	676	4	US-09-151-189-2	Sequence 2, Appl
134	5	17.2	467	2	US-08-819-825-3	Sequence 3, Appl	207	5	17.2	706	1	US-08-484-105-16	Sequence 16, Appl
135	5	17.2	467	4	US-09-155-855-3	Sequence 3, Appl	208	5	17.2	706	1	US-08-484-106-16	Sequence 16, Appl
136	5	17.2	467	4	US-09-163-642-3	Sequence 3, Appl	209	5	17.2	718	1	US-08-190-802A-44	Sequence 44, Appl
137	5	17.2	467	4	US-09-543-744-3	Sequence 3, Appl	210	5	17.2	718	4	US-08-477-346-44	Sequence 44, Appl
138	5	17.2	467	4	US-09-233-510-32	Sequence 32, Appl	211	5	17.2	718	4	US-08-473-089-44	Sequence 44, Appl
139	5	17.2	473	2	US-08-759-581B-15	Sequence 15, Appl	212	5	17.2	793	2	US-08-720-484A-2	Sequence 2, Appl
140	5	17.2	473	4	US-09-304-711-15	Sequence 15, Appl	213	5	17.2	793	4	US-08-953-823A-2	Sequence 2, Appl
141	5	17.2	492	4	US-08-845-258-21	Sequence 21, Appl	214	5	17.2	793	4	US-09-293-505-16	Sequence 16, Appl
142	5	17.2	492	4	US-08-990-571-21	Sequence 21, Appl	215	5	17.2	793	4	US-09-293-505-17	Sequence 17, Appl
143	5	17.2	492	4	US-08-723-142A-21	Sequence 21, Appl	216	5	17.2	803	4	US-09-293-505-15	Sequence 15, Appl
144	5	17.2	500	1	US-07-755-73C-8	Sequence 8, Appl	217	5	17.2	809	5	PCT-US91-01726-3	Sequence 3, Appl
145	5	17.2	521	1	US-08-584-226-2	Sequence 2, Appl	218	5	17.2	861	3	US-09-022-875-2	Sequence 2, Appl
146	5	17.2	525	4	US-08-542-634-28	Sequence 28, Appl	219	5	17.2	908	5	PCT-US95-03747-3	Sequence 3, Appl
147	5	17.2	525	5	PCT-US95-13703-28	Sequence 28, Appl	220	5	17.2	912	5	PCT-US95-03747-2	Sequence 2, Appl
148	5	17.2	526	3	US-08-673-814-3	Sequence 3, Appl	221	5	17.2	922	4	US-09-141-206-6	Sequence 6, Appl
149	5	17.2	526	4	US-09-115-824-3	Sequence 3, Appl	222	5	17.2	932	3	US-08-968-752B-6	Sequence 6, Appl
150	5	17.2	529	1	US-08-152-019A-40	Sequence 40, Appl	223	5	17.2	933	4	US-09-141-206-2	Sequence 2, Appl
151	5	17.2	529	3	US-08-650-599A-3	Sequence 3, Appl	224	5	17.2	933	4	US-09-107-149-2	Sequence 2, Appl
152	5	17.2	529	4	US-09-490-517-3	Sequence 3, Appl	225	5	17.2	933	4	US-09-107-149-17	Sequence 17, Appl
153	5	17.2	538	1	US-07-914-282D-5	Sequence 5, Appl	226	5	17.2	946	3	US-09-074-579-3	Sequence 3, Appl
154	5	17.2	538	1	US-08-276-887A-5	Sequence 5, Appl	227	5	17.2	946	4	US-09-388-774-3	Sequence 3, Appl
155	5	17.2	538	5	PCT-US93-02460-5	Sequence 5, Appl	228	5	17.2	950	4	US-09-449-285A-4	Sequence 4, Appl
156	5	17.2	540	2	US-08-884-072-3	Sequence 3, Appl	229	5	17.2	1002	4	US-09-268-347-24	Sequence 24, Appl
157	5	17.2	540	4	US-08-542-634-26	Sequence 26, Appl	230	5	17.2	1046	1	US-08-386-727-2	Sequence 2, Appl
158	5	17.2	540	4	US-09-212-168-3	Sequence 3, Appl	231	5	17.2	1046	2	US-08-600-452A-2	Sequence 2, Appl
159	5	17.2	540	5	PCT-US95-13703-26	Sequence 26, Appl	232	5	17.2	1085	1	US-08-485-588-5	Sequence 5, Appl
160	5	17.2	543	3	US-08-808-931-12	Sequence 12, Appl	233	5	17.2	1085	2	US-08-484-565-5	Sequence 5, Appl
161	5	17.2	543	3	US-08-808-323-12	Sequence 12, Appl	234	5	17.2	1085	2	US-08-480-751-5	Sequence 5, Appl
162	5	17.2	543	3	US-09-050-603A-12	Sequence 12, Appl	235	5	17.2	1085	2	US-08-943-986-5	Sequence 5, Appl
163	5	17.2	543	3	US-09-102-420B-12	Sequence 12, Appl	236	5	17.2	1085	3	US-08-353-784-5	Sequence 5, Appl
164	5	17.2	543	4	US-09-497-698-12	Sequence 12, Appl	237	5	17.2	1085	3	US-08-484-719B-5	Sequence 5, Appl
165	5	17.2	548	4	US-08-477-292-16	Sequence 16, Appl	238	5	17.2	1085	4	US-08-484-159-5	Sequence 5, Appl
166	5	17.2	549	4	US-08-542-634-16	Sequence 16, Appl	239	5	17.2	1091	3	US-08-986-485-5	Sequence 5, Appl
167	5	17.2	549	5	PCT-US95-13703-16	Sequence 16, Appl	240	5	17.2	1147	1	US-08-144-121-3	Sequence 3, Appl
168	5	17.2	557	4	US-09-413-814-34	Sequence 34, Appl	241	5	17.2	1147	2	US-08-735-893-3	Sequence 3, Appl
169	5	17.2	575	4	US-09-171-461-30	Sequence 30, Appl	242	5	17.2	1165	1	US-08-144-121-2	Sequence 2, Appl
170	5	17.2	576	1	US-08-221-817-13	Sequence 13, Appl	243	5	17.2	1165	2	US-08-735-893-2	Sequence 2, Appl
171	5	17.2	576	1	US-08-454-439-13	Sequence 13, Appl	244	5	17.2	1242	3	US-08-680-326-33	Sequence 33, Appl
172	5	17.2	576	4	US-08-464-954B-6	Sequence 6, Appl	245	5	17.2	1410	3	US-09-325-409-3	Sequence 3, Appl
173	5	17.2	576	5	PCT-US94-10487-13	Sequence 13, Appl	246	5	17.2	1410	4	US-09-368-102-3	Sequence 3, Appl

247	5	17.2	1410	4	US-09-567-969-3	Sequence 3, Appl	Sequence 3, Appl	320	4	13.8	10	3	US-09-231-797-17	Sequence 17, Appl
248	5	17.2	1410	4	US-09-568-480-3	Sequence 3, Appl	Sequence 3, Appl	321	4	13.8	10	3	US-08-934-224-17	Sequence 17, Appl
249	5	17.2	1410	4	US-09-568-486-3	Sequence 3, Appl	Sequence 3, Appl	322	4	13.8	10	3	US-09-105-678A-51	Sequence 51, Appl
250	5	17.2	1410	4	US-09-568-472-3	Sequence 3, Appl	Sequence 3, Appl	323	4	13.8	10	3	US-08-933-843-17	Sequence 17, Appl
251	5	17.2	1456	1	US-08-026-138E-8	Sequence 8, Appl	Sequence 8, Appl	324	4	13.8	10	3	US-09-244-298A-71	Sequence 71, Appl
252	5	17.2	1482	1	US-08-026-138E-2	Sequence 2, Appl	Sequence 2, Appl	325	4	13.8	10	4	US-08-934-223-17	Sequence 17, Appl
253	5	17.2	1484	2	US-08-231-193A-56	Sequence 56, Appl	Sequence 56, Appl	326	4	13.8	10	4	US-09-516-704-71	Sequence 71, Appl
254	5	17.2	1484	2	US-08-486-273A-56	Sequence 56, Appl	Sequence 56, Appl	327	4	13.8	10	4	US-09-413-492-17	Sequence 17, Appl
255	5	17.2	1484	3	US-08-940-086A-56	Sequence 56, Appl	Sequence 56, Appl	328	4	13.8	10	4	US-09-421-208-51	Sequence 51, Appl
256	5	17.2	1484	4	US-08-940-035A-56	Sequence 56, Appl	Sequence 56, Appl	329	4	13.8	10	5	PCT-US91-07506-62	Sequence 62, Appl
257	5	17.2	1525	3	US-09-131-847-2	Sequence 2, Appl	Sequence 2, Appl	330	4	13.8	10	5	PCT-US96-01735-4	Sequence 4, Appl
258	5	17.2	1525	4	US-09-540-245A-2	Sequence 2, Appl	Sequence 2, Appl	331	4	13.8	11	1	US-08-594-447-34	Sequence 34, Appl
259	5	17.2	1525	4	US-09-540-153-2	Sequence 2, Appl	Sequence 2, Appl	332	4	13.8	11	1	US-08-541-964-33	Sequence 33, Appl
260	5	17.2	1572	3	US-08-290-731C-5	Sequence 5, Appl	Sequence 5, Appl	333	4	13.8	11	2	US-08-835-231-2	Sequence 2, Appl
261	5	17.2	1596	3	US-09-356-952-3	Sequence 3, Appl	Sequence 3, Appl	334	4	13.8	11	2	US-08-665-647-48	Sequence 48, Appl
262	5	17.2	1713	3	US-08-600-982-24	Sequence 24, Appl	Sequence 24, Appl	335	4	13.8	11	3	US-08-968-747-11	Sequence 11, Appl
263	5	17.2	1713	5	PCT-US94-10261A-24	Sequence 24, Appl	Sequence 24, Appl	336	4	13.8	11	3	US-08-893-526A-16	Sequence 16, Appl
264	5	17.2	1780	1	US-08-769-309A-5	Sequence 5, Appl	Sequence 5, Appl	337	4	13.8	11	3	US-08-493-071-27	Sequence 27, Appl
265	5	17.2	1780	3	US-08-994-570-5	Sequence 5, Appl	Sequence 5, Appl	338	4	13.8	11	4	US-09-417-305-4	Sequence 4, Appl
266	5	17.2	1956	4	US-08-843-417-10	Sequence 10, Appl	Sequence 10, Appl	339	4	13.8	11	4	US-09-108-661-2	Sequence 2, Appl
267	5	17.2	2016	4	US-09-634-920-4	Sequence 4, Appl	Sequence 4, Appl	340	4	13.8	12	6	5284931-36	Patent No. 5284931
268	5	17.2	2584	3	US-08-936-135-4	Sequence 4, Appl	Sequence 4, Appl	341	4	13.8	12	6	5451527-7	Patent No. 5451527
269	5	17.2	15281	2	US-08-471-119A-2	Sequence 2, Appl	Sequence 2, Appl	342	4	13.8	13	1	US-08-032-842-19	Sequence 19, Appl
270	4	13.8	4	2	US-08-478-386A-11	Sequence 11, Appl	Sequence 11, Appl	343	4	13.8	13	2	US-08-694-579-9	Sequence 9, Appl
271	4	13.8	4	2	US-08-478-386A-13	Sequence 13, Appl	Sequence 13, Appl	344	4	13.8	13	2	US-08-948-155-9	Sequence 9, Appl
272	4	13.8	4	2	US-08-292-597-11	Sequence 11, Appl	Sequence 11, Appl	345	4	13.8	14	1	US-08-471-033-12	Sequence 12, Appl
273	4	13.8	4	2	US-08-292-597-13	Sequence 13, Appl	Sequence 13, Appl	346	4	13.8	14	1	US-08-615-279-31	Sequence 31, Appl
274	4	13.8	4	2	US-08-388-653-11	Sequence 11, Appl	Sequence 11, Appl	347	4	13.8	14	2	US-08-471-044-12	Sequence 12, Appl
275	4	13.8	4	2	US-08-388-653-13	Sequence 13, Appl	Sequence 13, Appl	348	4	13.8	14	2	US-08-463-483A-12	Sequence 12, Appl
276	4	13.8	4	2	US-08-473-985-11	Sequence 11, Appl	Sequence 11, Appl	349	4	13.8	14	2	US-08-471-046A-12	Sequence 12, Appl
277	4	13.8	4	2	US-08-473-985-13	Sequence 13, Appl	Sequence 13, Appl	350	4	13.8	14	2	US-08-470-566B-12	Sequence 12, Appl
278	4	13.8	4	2	US-08-483-898-11	Sequence 11, Appl	Sequence 11, Appl	351	4	13.8	14	2	US-08-838-219B-12	Sequence 12, Appl
279	4	13.8	4	2	US-08-483-898-13	Sequence 13, Appl	Sequence 13, Appl	352	4	13.8	14	2	US-08-765-452-10	Sequence 10, Appl
280	4	13.8	4	3	US-09-087-716-11	Sequence 11, Appl	Sequence 11, Appl	353	4	13.8	14	2	US-08-469-334-12	Sequence 12, Appl
281	4	13.8	4	3	US-09-087-716-13	Sequence 13, Appl	Sequence 13, Appl	354	4	13.8	14	3	US-09-300-529-12	Sequence 12, Appl
282	4	13.8	4	3	US-09-157-753-11	Sequence 11, Appl	Sequence 11, Appl	355	4	13.8	14	3	US-09-233-336A-12	Sequence 12, Appl
283	4	13.8	4	3	US-09-157-753-13	Sequence 13, Appl	Sequence 13, Appl	356	4	13.8	14	4	US-09-233-752A-12	Sequence 12, Appl
284	4	13.8	4	3	US-09-157-230-11	Sequence 11, Appl	Sequence 11, Appl	357	4	13.8	14	4	US-09-402-036-12	Sequence 12, Appl
285	4	13.8	4	3	US-09-157-230-13	Sequence 13, Appl	Sequence 13, Appl	358	4	13.8	14	4	US-08-481-968A-18	Sequence 18, Appl
286	4	13.8	4	3	US-09-087-811-11	Sequence 11, Appl	Sequence 11, Appl	359	4	13.8	14	4	US-08-154-712B-18	Sequence 18, Appl
287	4	13.8	4	3	US-09-087-811-13	Sequence 13, Appl	Sequence 13, Appl	360	4	13.8	15	1	US-08-222-619-9	Sequence 9, Appl
288	4	13.8	4	3	US-09-156-855-11	Sequence 11, Appl	Sequence 11, Appl	361	4	13.8	15	1	US-08-277-299-2	Sequence 2, Appl
289	4	13.8	4	3	US-09-156-855-13	Sequence 13, Appl	Sequence 13, Appl	362	4	13.8	15	1	US-08-737-927-2	Sequence 2, Appl
290	4	13.8	4	4	US-09-158-010-11	Sequence 11, Appl	Sequence 11, Appl	363	4	13.8	15	2	US-08-880-671-2	Sequence 2, Appl
291	4	13.8	4	4	US-09-158-010-13	Sequence 13, Appl	Sequence 13, Appl	364	4	13.8	15	2	US-08-986-234-7	Sequence 7, Appl
292	4	13.8	4	4	US-09-087-647-11	Sequence 11, Appl	Sequence 11, Appl	365	4	13.8	15	5	US-09-177-249-228	Sequence 228, Appl
293	4	13.8	4	4	US-09-087-647-13	Sequence 13, Appl	Sequence 13, Appl	366	4	13.8	15	5	PCT-US95-04075-9	Sequence 9, Appl
294	4	13.8	4	4	US-09-180-903-2	Sequence 2, Appl	Sequence 2, Appl	367	4	13.8	16	1	US-07-942-245-98	Sequence 98, Appl
295	4	13.8	4	4	US-09-302-629-11	Sequence 11, Appl	Sequence 11, Appl	368	4	13.8	16	1	US-08-222-619-30	Sequence 30, Appl
296	4	13.8	4	4	US-09-302-629-13	Sequence 13, Appl	Sequence 13, Appl	369	4	13.8	16	1	US-08-307-724B-23	Sequence 23, Appl
297	4	13.8	5	1	US-07-695-472B-8	Sequence 8, Appl	Sequence 8, Appl	370	4	13.8	16	4	US-08-602-999A-406	Sequence 406, Appl
298	4	13.8	6	4	US-08-187-859-3069	Sequence 3069, Ap	Sequence 3069, Ap	371	4	13.8	16	4	US-09-031-295-4	Sequence 4, Appl
299	4	13.8	7	2	US-08-595-043A-71	Sequence 71, Appl	Sequence 71, Appl	372	4	13.8	16	5	PCT-US95-04075-30	Sequence 30, Appl
300	4	13.8	7	4	US-09-187-859-3070	Sequence 3070, Ap	Sequence 3070, Ap	373	4	13.8	16	6	5449751-4	Patent No. 5449751
301	4	13.8	7	6	5284931-27	Patent No. 5284931	Patent No. 5284931	374	4	13.8	17	1	US-08-664-449-35	Sequence 35, Appl
302	4	13.8	8	1	US-08-206-310A-1	Sequence 1, Appl	Sequence 1, Appl	375	4	13.8	17	2	US-08-871-074-7	Sequence 7, Appl
303	4	13.8	8	2	US-08-398-010A-1	Sequence 1, Appl	Sequence 1, Appl	376	4	13.8	17	2	US-08-871-074-18	Sequence 18, Appl
304	4	13.8	8	2	US-08-398-628A-1	Sequence 1, Appl	Sequence 1, Appl	377	4	13.8	17	3	US-08-819-177-22	Sequence 22, Appl
305	4	13.8	8	2	US-08-399-115A-1	Sequence 1, Appl	Sequence 1, Appl	378	4	13.8	17	5	PCT-US94-05471-7	Sequence 7, Appl
306	4	13.8	8	4	US-09-187-859-3071	Sequence 3071, Ap	Sequence 3071, Ap	379	4	13.8	17	5	PCT-US94-05471-18	Sequence 18, Appl
307	4	13.8	9	2	US-08-340-283-34	Sequence 34, Appl	Sequence 34, Appl	380	4	13.8	18	2	US-08-480-190-242	Sequence 242, Appl
308	4	13.8	9	2	US-08-340-283-167	Sequence 167, App	Sequence 167, App	381	4	13.8	18	2	US-08-488-379-242	Sequence 242, Appl
309	4	13.8	9	2	US-08-340-283-194	Sequence 194, App	Sequence 194, App	382	4	13.8	18	2	US-08-871-074-19	Sequence 19, Appl
310	4	13.8	10	1	US-08-346-333-62	Sequence 62, Appl	Sequence 62, Appl	383	4	13.8	18	3	US-08-940-095-237	Sequence 237, Appl
311	4	13.8	10	2	US-08-248-839C-180	Sequence 180, App	Sequence 180, App	384	4	13.8	18	3	US-08-940-096-237	Sequence 237, Appl
312	4	13.8	10	2	US-08-764-640-71	Sequence 71, Appl	Sequence 71, Appl	385	4	13.8	18	3	US-08-940-096-237	Sequence 237, Appl
313	4	13.8	10	2	US-08-934-222-17	Sequence 17, Appl	Sequence 17, Appl	386	4	13.8	18	4	US-09-184-938-10	Sequence 10, Appl
314	4	13.8	10	2	US-08-933-402-17	Sequence 17, Appl	Sequence 17, Appl	387	4	13.8	18	4	US-09-465-719-237	Sequence 237, Appl
315	4	13.8	10	2	US-09-207-621-17	Sequence 17, Appl	Sequence 17, Appl	388	4	13.8	18	4	US-09-453-605-237	Sequence 237, Appl
316	4	13.8	10	2	US-08-532-818-17	Sequence 17, Appl	Sequence 17, Appl	389	4	13.8	18	5	PCT-US93-07545-242	Sequence 242, Appl
317	4	13.8	10	3	US-08-159-339A-79	Sequence 79, Appl	Sequence 79, Appl	390	4	13.8	18	5	PCT-US94-05471-19	Sequence 19, Appl
318	4	13.8	10	3	US-08-973-225-71	Sequence 71, Appl	Sequence 71, Appl	391	4	13.8	19	3	US-08-974-022-40	Sequence 40, Appl
319	4	13.8	10	3	US-08-973-225-206	Sequence 206, App	Sequence 206, App	392	4	13.8	19	4	US-08-795-445A-40	Sequence 40, Appl

393	4	13.8	19	4	US-08-795-447A-40	Sequence 40, Appl	466	4	13.8	32	1	US-08-471-052A-17	Sequence 17, Appl
394	4	13.8	19	4	US-08-974-186-40	Sequence 40, Appl	467	4	13.8	32	1	US-08-451-240-2	Sequence 2, Appl
395	4	13.8	19	4	US-08-795-446B-40	Sequence 40, Appl	468	4	13.8	32	1	US-08-451-240-5	Sequence 5, Appl
396	4	13.8	20	1	US-08-484-135-67	Sequence 67, Appl	469	4	13.8	32	1	US-08-451-240-15	Sequence 15, Appl
397	4	13.8	20	1	US-08-484-635-23	Sequence 23, Appl	470	4	13.8	32	1	US-08-486-013-19	Sequence 19, Appl
398	4	13.8	20	2	US-08-484-631-23	Sequence 23, Appl	471	4	13.8	32	1	US-08-189-331-17	Sequence 17, Appl
399	4	13.8	20	2	US-08-934-915-165	Sequence 165, App	472	4	13.8	32	1	US-08-737-927-4	Sequence 4, Appl
400	4	13.8	20	2	US-08-827-570-23	Sequence 23, Appl	473	4	13.8	32	1	US-08-482-279-19	Sequence 19, Appl
401	4	13.8	20	3	US-09-100-409A-25	Sequence 25, Appl	474	4	13.8	32	2	US-08-342-268-19	Sequence 19, Appl
402	4	13.8	20	3	US-08-825-852-60	Sequence 60, Appl	475	4	13.8	32	2	US-08-471-939-17	Sequence 17, Appl
403	4	13.8	20	4	US-09-052-888-61	Sequence 61, Appl	476	4	13.8	32	2	US-08-470-846A-5	Sequence 5, Appl
404	4	13.8	21	1	US-07-746-705A-8	Sequence 8, Appl	477	4	13.8	32	2	US-08-470-846A-15	Sequence 15, Appl
405	4	13.8	21	1	US-08-447-925-4	Sequence 4, Appl	478	4	13.8	32	2	US-08-470-846A-18	Sequence 18, Appl
406	4	13.8	21	2	US-08-380-182-7	Sequence 7, Appl	479	4	13.8	32	2	US-08-471-800-17	Sequence 17, Appl
407	4	13.8	21	2	US-08-637-759B-485	Sequence 485, App	480	4	13.8	32	2	US-08-471-068-17	Sequence 17, Appl
408	4	13.8	21	3	US-08-871-355A-485	Sequence 485, App	481	4	13.8	32	3	US-09-015-968-19	Sequence 19, Appl
409	4	13.8	21	4	US-08-423-646A-46	Sequence 46, Appl	482	4	13.8	32	4	US-09-068-738A-2	Sequence 2, Appl
410	4	13.8	21	4	US-09-201-945-485	Sequence 485, App	483	4	13.8	32	4	US-09-068-738A-17	Sequence 17, Appl
411	4	13.8	23	1	US-07-946-497-8	Sequence 8, Appl	484	4	13.8	32	4	US-09-397-386-19	Sequence 19, Appl
412	4	13.8	23	1	US-08-483-322-8	Sequence 8, Appl	485	4	13.8	32	5	PCT-US94-12591-2	Sequence 2, Appl
413	4	13.8	23	2	US-08-478-882-8	Sequence 8, Appl	486	4	13.8	32	5	PCT-US94-12591-5	Sequence 5, Appl
414	4	13.8	23	4	US-08-882-046-10	Sequence 10, Appl	487	4	13.8	32	5	PCT-US94-12591-15	Sequence 15, Appl
415	4	13.8	24	1	US-08-018-129-19	Sequence 19, Appl	488	4	13.8	32	6	5449751-1	Patent No. 5449751
416	4	13.8	24	1	US-08-220-272A-6	Sequence 6, Appl	489	4	13.8	32	6	5449751-3	Patent No. 5449751
417	4	13.8	24	2	US-08-448-250-19	Sequence 19, Appl	490	4	13.8	33	1	US-08-438-753B-8	Sequence 8, Appl
418	4	13.8	24	3	US-08-630-916A-110	Sequence 110, App	491	4	13.8	33	1	US-08-443-883A-8	Sequence 8, Appl
419	4	13.8	24	4	US-08-913-942-9	Sequence 9, Appl	492	4	13.8	33	1	US-08-309-512-12	Sequence 12, Appl
420	4	13.8	24	4	US-08-913-942-10	Sequence 10, Appl	493	4	13.8	33	2	US-08-461-990B-7	Sequence 7, Appl
421	4	13.8	24	4	US-08-913-942-11	Sequence 11, Appl	494	4	13.8	33	2	US-08-631-328-8	Sequence 8, Appl
422	4	13.8	24	4	US-08-913-942-13	Sequence 13, Appl	495	4	13.8	33	2	US-08-455-24B-8	Sequence 8, Appl
423	4	13.8	24	4	US-08-630-915A-166	Sequence 166, App	496	4	13.8	33	2	US-08-455-021B-8	Sequence 8, Appl
424	4	13.8	24	4	US-09-315-304B-1522	Sequence 1522, Ap	497	4	13.8	33	4	US-09-045-467-8	Sequence 8, Appl
425	4	13.8	25	2	US-08-726-306A-169	Sequence 169, App	498	4	13.8	34	1	US-08-158-189-28	Sequence 28, Appl
426	4	13.8	26	2	US-08-620-151-41	Sequence 41, Appl	499	4	13.8	34	2	US-08-491-204A-12	Sequence 12, Appl
427	4	13.8	26	2	US-08-563-892A-28	Sequence 28, Appl	500	4	13.8	34	4	US-09-314-268-158	Sequence 158, App
428	4	13.8	27	2	US-07-684-965-8	Sequence 8, Appl	501	4	13.8	35	3	US-09-100-600A-1	Sequence 1, Appl
429	4	13.8	27	2	US-08-819-177-23	Sequence 23, Appl	502	4	13.8	35	3	US-09-100-600A-2	Sequence 2, Appl
430	4	13.8	27	4	US-09-252-658-8	Sequence 8, Appl	503	4	13.8	35	4	US-09-082-279B-646	Sequence 646, App
431	4	13.8	27	4	US-09-082-279B-1086	Sequence 1086, Ap	504	4	13.8	35	4	US-09-082-279B-647	Sequence 647, App
432	4	13.8	28	3	US-09-315-304B-1086	Sequence 1086, Ap	505	4	13.8	35	4	US-09-082-279B-648	Sequence 648, App
433	4	13.8	28	3	US-09-045-632-95	Sequence 95, Appl	506	4	13.8	35	4	US-09-082-279B-649	Sequence 649, App
434	4	13.8	28	3	US-08-486-099-86	Sequence 86, Appl	507	4	13.8	35	4	US-09-082-279B-650	Sequence 650, App
435	4	13.8	28	3	US-08-916-043-4	Sequence 4, Appl	508	4	13.8	35	4	US-09-082-279B-651	Sequence 651, App
436	4	13.8	28	3	US-08-360-107A-96	Sequence 96, Appl	509	4	13.8	35	4	US-09-315-304B-646	Sequence 646, App
437	4	13.8	28	3	US-08-484-223B-86	Sequence 86, Appl	510	4	13.8	35	4	US-09-315-304B-647	Sequence 647, App
438	4	13.8	28	3	US-08-919-597-86	Sequence 86, Appl	511	4	13.8	35	4	US-09-315-304B-648	Sequence 648, App
439	4	13.8	28	3	US-08-475-668A-86	Sequence 86, Appl	512	4	13.8	35	4	US-09-315-304B-649	Sequence 649, App
440	4	13.8	28	3	US-08-485-551A-86	Sequence 86, Appl	513	4	13.8	35	4	US-09-315-304B-650	Sequence 650, App
441	4	13.8	28	3	US-08-471-913A-86	Sequence 86, Appl	514	4	13.8	35	4	US-09-315-304B-651	Sequence 651, App
442	4	13.8	28	4	US-08-485-264A-86	Sequence 86, Appl	515	4	13.8	36	1	US-08-118-270-231	Sequence 231, App
443	4	13.8	28	4	US-08-474-349A-86	Sequence 86, Appl	516	4	13.8	36	1	US-08-487-890A-13	Sequence 13, Appl
444	4	13.8	29	4	US-09-461-697-435	Sequence 435, App	517	4	13.8	36	2	US-08-478-435-13	Sequence 13, Appl
445	4	13.8	30	1	US-08-158-189-51	Sequence 51, Appl	518	4	13.8	36	2	US-08-337-483-13	Sequence 13, Appl
446	4	13.8	30	2	US-08-491-204A-6	Sequence 6, Appl	519	4	13.8	36	2	US-08-478-373-13	Sequence 13, Appl
447	4	13.8	30	2	US-08-310-912A-153	Sequence 153, App	520	4	13.8	36	3	US-08-474-671-13	Sequence 13, Appl
448	4	13.8	30	4	US-09-301-085-153	Sequence 153, App	521	4	13.8	36	3	US-08-483-577A-13	Sequence 13, Appl
449	4	13.8	30	5	PCT-US95-04589-153	Sequence 153, App	522	4	13.8	36	4	US-08-897-438-13	Sequence 13, Appl
450	4	13.8	31	1	US-08-190-802A-98	Sequence 98, Appl	523	4	13.8	36	4	US-08-637-654-13	Sequence 13, Appl
451	4	13.8	31	1	US-08-418-716A-1	Sequence 1, Appl	524	4	13.8	36	5	PCT-US93-08528-231	Sequence 231, App
452	4	13.8	31	1	US-08-418-716A-2	Sequence 2, Appl	525	4	13.8	37	1	US-08-212-236-10	Sequence 10, Appl
453	4	13.8	31	1	US-08-418-716A-4	Sequence 4, Appl	526	4	13.8	37	1	US-08-477-727A-102	Sequence 102, App
454	4	13.8	31	1	US-08-418-716A-5	Sequence 5, Appl	527	4	13.8	37	2	US-08-471-675A-24	Sequence 24, Appl
455	4	13.8	31	1	US-08-418-716A-6	Sequence 6, Appl	528	4	13.8	37	2	US-08-892-549-6	Sequence 6, Appl
456	4	13.8	31	1	US-08-418-716A-7	Sequence 7, Appl	529	4	13.8	37	3	US-08-892-549-28	Sequence 28, Appl
457	4	13.8	31	4	US-08-602-999A-54	Sequence 54, Appl	530	4	13.8	37	3	US-08-819-177-24	Sequence 24, Appl
458	4	13.8	31	4	US-08-477-346-98	Sequence 98, Appl	531	4	13.8	37	3	US-08-302-069A-23	Sequence 23, Appl
459	4	13.8	31	4	US-09-070-504-23	Sequence 23, Appl	532	4	13.8	37	4	US-08-751-344B-16	Sequence 16, Appl
460	4	13.8	31	4	US-08-278-865-54	Sequence 54, Appl	533	4	13.8	37	4	US-08-751-344B-19	Sequence 19, Appl
461	4	13.8	31	4	US-09-011-922A-3	Sequence 3, Appl	534	4	13.8	38	4	US-09-324-455-16	Sequence 16, Appl
462	4	13.8	31	4	US-09-011-922A-14	Sequence 14, Appl	535	4	13.8	38	4	US-09-172-841-9	Sequence 9, Appl
463	4	13.8	31	4	US-09-486-580A-6	Sequence 6, Appl	536	4	13.8	38	4	US-09-172-841-11	Sequence 11, Appl
464	4	13.8	31	4	US-08-473-089-98	Sequence 98, Appl	537	4	13.8	39	1	US-07-930-649-14	Sequence 14, Appl
465	4	13.8	32	1	US-08-176-500-17	Sequence 17, Appl	538	4	13.8	39	1	US-08-486-013-11	Sequence 11, Appl

539	4	13.8	39	2	US-08-482-279-11	Sequence 11, Appl	612	4	13.8	53	2	US-08-563-892A-29	Sequence 29, Appl
540	4	13.8	39	2	US-08-342-268-11	Sequence 11, Appl	613	4	13.8	54	4	US-09-227-357-150	Sequence 150, App
541	4	13.8	39	3	US-09-015-968-11	Sequence 11, Appl	614	4	13.8	55	3	US-09-320-095-8	Sequence 8, Appli
542	4	13.8	39	4	US-09-397-386-11	Sequence 11, Appl	615	4	13.8	55	4	US-08-349-403-5	Sequence 5, Appli
543	4	13.8	39	4	US-08-860-089-1	Sequence 1, Appli	616	4	13.8	55	4	US-09-523-487-8	Sequence 8, Appli
544	4	13.8	39	5	PCT-US93-05235-14	Sequence 14, Appl	617	4	13.8	56	1	US-08-118-270-250	Sequence 250, App
545	4	13.8	40	1	US-08-033-873-10	Sequence 10, Appl	618	4	13.8	56	5	PCT-US93-08528-250	Sequence 250, App
546	4	13.8	40	1	US-08-467-943-1	Sequence 1, Appli	619	4	13.8	56	5	PCT-US94-06655-3	Sequence 3, Appli
547	4	13.8	40	2	US-08-356-832-10	Sequence 10, Appl	620	4	13.8	57	1	US-08-055-988-2	Sequence 2, Appli
548	4	13.8	40	4	US-08-988-705-10	Sequence 10, Appl	621	4	13.8	57	1	US-08-118-270-251	Sequence 251, App
549	4	13.8	40	4	US-08-936-165A-353	Sequence 353, App	622	4	13.8	57	1	US-08-358-160-102	Sequence 102, App
550	4	13.8	41	1	US-08-486-013-15	Sequence 15, Appl	623	4	13.8	57	1	US-08-358-160-103	Sequence 103, App
551	4	13.8	41	2	US-08-482-279-15	Sequence 15, Appl	624	4	13.8	57	1	US-08-358-160-104	Sequence 104, App
552	4	13.8	41	3	US-08-342-268-15	Sequence 15, Appl	625	4	13.8	57	1	US-08-358-160-156	Sequence 156, App
553	4	13.8	41	3	US-09-015-968-15	Sequence 15, Appl	626	4	13.8	57	4	US-09-227-357-275	Sequence 275, App
554	4	13.8	41	4	US-09-397-386-15	Sequence 15, Appl	627	4	13.8	57	5	PCT-US93-08528-251	Sequence 251, App
555	4	13.8	41	6	517197-45	Patent No. 517197	628	4	13.8	57	6	5466783-17	Patent No. 5466783
556	4	13.8	42	1	US-08-447-925-2	Sequence 2, Appli	629	4	13.8	58	1	US-08-384-489-16	Sequence 16, Appl
557	4	13.8	42	1	US-08-099-354-8	Sequence 8, Appli	630	4	13.8	58	1	US-08-334-773A-1	Sequence 1, Appli
558	4	13.8	42	1	US-08-664-449-39	Sequence 39, Appl	631	4	13.8	58	1	US-08-334-773A-2	Sequence 2, Appli
559	4	13.8	42	2	US-08-288-059-34	Sequence 34, Appl	632	4	13.8	58	1	US-08-358-160-79	Sequence 79, Appl
560	4	13.8	42	2	US-07-801-798-3	Sequence 3, Appli	633	4	13.8	58	1	US-08-463-155A-40	Sequence 40, Appl
561	4	13.8	42	4	US-08-988-856B-31	Sequence 31, Appl	634	4	13.8	58	1	US-08-463-432B-40	Sequence 40, Appl
562	4	13.8	42	5	PCT-US92-10140-3	Sequence 3, Appli	635	4	13.8	58	1	US-08-676-125A-37	Sequence 37, Appl
563	4	13.8	43	2	US-08-659-251-34	Sequence 34, Appl	636	4	13.8	58	1	US-08-676-125A-38	Sequence 38, Appl
564	4	13.8	43	2	US-08-659-251-35	Sequence 35, Appl	637	4	13.8	58	1	US-08-206-310A-40	Sequence 40, Appl
565	4	13.8	43	2	US-08-659-251-36	Sequence 36, Appl	638	4	13.8	58	1	US-08-398-010A-40	Sequence 40, Appl
566	4	13.8	43	4	US-09-417-305-1	Sequence 1, Appli	639	4	13.8	58	2	US-08-398-628A-40	Sequence 40, Appl
567	4	13.8	43	4	US-08-905-223-293	Sequence 293, App	640	4	13.8	58	2	US-08-399-115A-40	Sequence 40, Appl
568	4	13.8	43	4	US-09-256-490-34	Sequence 34, Appl	641	4	13.8	58	2	US-09-136-012A-37	Sequence 37, Appl
569	4	13.8	43	4	US-09-256-490-35	Sequence 35, Appl	642	4	13.8	58	2	US-09-136-012A-38	Sequence 38, Appl
570	4	13.8	43	4	US-09-256-490-36	Sequence 36, Appl	643	4	13.8	58	3	US-08-676-124-68	Sequence 68, Appl
571	4	13.8	43	4	US-09-227-357-614	Sequence 614, App	644	4	13.8	58	3	US-08-676-124-69	Sequence 69, Appl
572	4	13.8	43	4	US-09-314-268-157	Sequence 157, App	645	4	13.8	58	3	US-09-414-878-68	Sequence 68, Appl
573	4	13.8	43	5	PCT-US96-11445-34	Sequence 34, Appl	646	4	13.8	58	3	US-09-414-878-69	Sequence 69, Appl
574	4	13.8	43	5	PCT-US96-11445-35	Sequence 35, Appl	647	4	13.8	58	3	US-09-240-136-69	Sequence 69, Appl
575	4	13.8	43	5	PCT-US96-11445-36	Sequence 36, Appl	648	4	13.8	58	3	US-09-240-136-69	Sequence 69, Appl
576	4	13.8	44	1	US-08-262-037-130	Sequence 130, App	649	4	13.8	58	4	US-09-369-494-4	Sequence 4, Appli
577	4	13.8	44	1	US-08-486-013-17	Sequence 17, Appl	650	4	13.8	58	4	US-09-358-569D-4	Sequence 4, Appli
578	4	13.8	45	2	US-08-482-279-17	Sequence 17, Appl	651	4	13.8	58	6	5466783-15	Patent No. 5466783
579	4	13.8	45	2	US-08-342-268-17	Sequence 17, Appl	652	4	13.8	59	1	US-08-358-160-106	Sequence 106, App
580	4	13.8	45	2	US-08-805-191-3	Sequence 3, Appli	653	4	13.8	60	1	US-08-118-270-252	Sequence 252, App
581	4	13.8	45	2	US-09-015-968-17	Sequence 17, Appl	654	4	13.8	60	1	US-08-118-270-253	Sequence 253, App
582	4	13.8	45	4	US-08-905-223-318	Sequence 318, App	655	4	13.8	60	1	US-08-447-925-1	Sequence 1, Appli
583	4	13.8	45	4	US-08-905-223-367	Sequence 367, App	656	4	13.8	60	1	US-08-099-354-7	Sequence 7, Appli
584	4	13.8	45	4	US-08-975-080-18	Sequence 18, Appl	657	4	13.8	60	2	US-08-288-059-33	Sequence 33, Appl
585	4	13.8	45	4	US-09-397-386-17	Sequence 17, Appl	658	4	13.8	60	2	US-07-684-965-14	Sequence 14, Appl
586	4	13.8	45	3	US-09-100-600A-3	Sequence 3, Appli	659	4	13.8	60	4	US-09-252-658-14	Sequence 14, Appl
587	4	13.8	46	3	US-09-100-600A-38	Sequence 38, Appl	660	4	13.8	60	5	PCT-US93-08528-252	Sequence 252, App
588	4	13.8	46	4	US-09-124-671-5	Sequence 5, Appli	661	4	13.8	60	5	PCT-US93-08528-253	Sequence 253, App
589	4	13.8	46	4	US-08-679-493A-172	Sequence 172, App	662	4	13.8	62	2	US-08-563-892A-27	Sequence 27, Appl
590	4	13.8	47	3	US-08-753-007A-11	Sequence 11, Appl	663	4	13.8	65	3	US-09-100-600A-27	Sequence 27, Appl
591	4	13.8	47	4	US-09-398-496-11	Sequence 11, Appl	664	4	13.8	65	4	US-09-461-697-447	Sequence 447, App
592	4	13.8	48	1	US-08-625-322-21	Sequence 21, Appl	665	4	13.8	66	4	US-08-961-083-92	Sequence 92, Appl
593	4	13.8	48	3	US-08-974-022-43	Sequence 43, Appl	666	4	13.8	67	2	US-08-511-485-23	Sequence 23, Appl
594	4	13.8	48	4	US-08-795-445A-43	Sequence 43, Appl	667	4	13.8	67	4	US-09-461-697-445	Sequence 445, App
595	4	13.8	48	4	US-08-795-447A-43	Sequence 43, Appl	668	4	13.8	68	3	US-08-819-177-18	Sequence 18, Appl
596	4	13.8	48	4	US-08-974-186-43	Sequence 43, Appl	669	4	13.8	68	4	US-09-220-528-50	Sequence 50, Appl
597	4	13.8	48	4	US-08-795-446B-43	Sequence 43, Appl	670	4	13.8	68	4	US-09-220-528-81	Sequence 81, Appl
598	4	13.8	48	2	US-08-256-790-12	Sequence 12, Appl	671	4	13.8	69	4	US-08-855-526B-44	Sequence 44, Appl
599	4	13.8	50	2	US-08-142-551B-119	Sequence 119, App	672	4	13.8	69	5	PCT-US95-06406A-7	Sequence 7, Appli
600	4	13.8	50	4	US-09-171-646-1	Sequence 1, Appli	673	4	13.8	70	3	US-08-974-022-41	Sequence 41, Appl
601	4	13.8	51	1	US-08-418-893D-3	Sequence 3, Appli	674	4	13.8	70	4	US-08-795-445A-41	Sequence 41, Appl
602	4	13.8	51	3	US-09-044-536A-7	Sequence 7, Appli	675	4	13.8	70	4	US-08-795-447A-41	Sequence 41, Appl
603	4	13.8	51	4	US-09-369-494-10	Sequence 10, Appl	676	4	13.8	70	4	US-08-974-186-41	Sequence 41, Appl
604	4	13.8	52	1	US-08-118-270-237	Sequence 237, App	677	4	13.8	70	4	US-08-795-446B-41	Sequence 41, Appl
605	4	13.8	52	1	US-08-118-270-238	Sequence 238, App	678	4	13.8	70	6	5318899-16	Patent No. 5318899
606	4	13.8	52	2	US-08-856-444-3	Sequence 3, Appli	679	4	13.8	72	6	5318899-14	Patent No. 5318899
607	4	13.8	52	4	PCT-US93-08528-237	Sequence 14, Appl	680	4	13.8	75	3	US-09-100-600A-28	Sequence 28, Appl
608	4	13.8	52	5	PCT-US93-08528-237	Sequence 237, App	681	4	13.8	75	3	US-09-100-600A-29	Sequence 29, Appl
609	4	13.8	52	5	PCT-US93-08528-238	Sequence 238, App	682	4	13.8	76	1	US-08-667-679-3	Sequence 3, Appli
610	4	13.8	53	2	US-08-736-334B-7	Sequence 7, Appli	683	4	13.8	76	2	US-08-892-150-3	Sequence 3, Appli
611	4	13.8	53	2	US-08-256-790-8	Sequence 8, Appli	684	4	13.8	76	3	US-08-819-177-15	Sequence 15, Appl

685	4	13.8	77	1	US-07-689-693B-7	Sequence 7, Appli	758	4	13.8	101	1	US-08-340-203A-11	Sequence 11, Appli
686	4	13.8	77	2	US-08-716-308-17	Sequence 17, Appli	759	4	13.8	101	2	US-08-638-610-2	Sequence 2, Appli
687	4	13.8	78	1	US-08-491-357-4	Sequence 4, Appli	760	4	13.8	101	2	US-08-429-965-2	Sequence 2, Appli
688	4	13.8	78	2	US-08-612-788-20	Sequence 20, Appli	761	4	13.8	101	2	US-08-452-427-11	Sequence 11, Appli
689	4	13.8	78	3	US-09-066-028-20	Sequence 20, Appli	762	4	13.8	101	2	US-08-959-584-2	Sequence 2, Appli
690	4	13.8	78	3	US-08-968-633-4	Sequence 4, Appli	763	4	13.8	101	2	US-08-970-179A-2	Sequence 2, Appli
691	4	13.8	78	3	US-09-196-466-4	Sequence 4, Appli	764	4	13.8	101	2	US-08-969-249A-2	Sequence 2, Appli
692	4	13.8	78	5	PCT-US96-10823-4	Sequence 4, Appli	765	4	13.8	101	3	US-08-969-378-2	Sequence 2, Appli
693	4	13.8	79	4	US-08-432-994A-6	Sequence 6, Appli	766	4	13.8	101	3	US-09-085-407-11	Sequence 11, Appli
694	4	13.8	79	4	US-09-461-697-433	Sequence 433, App	767	4	13.8	101	3	US-09-136-594-2	Sequence 2, Appli
695	4	13.8	79	6	5284931-7	Patent No. 5284931	768	4	13.8	101	4	US-09-137-134-2	Sequence 2, Appli
696	4	13.8	81	1	US-08-123-702-8	Sequence 8, Appli	769	4	13.8	102	4	US-09-461-697-441	Sequence 441, App
697	4	13.8	81	4	US-09-247-155-125	Sequence 125, App	770	4	13.8	103	1	US-08-486-013-20	Sequence 20, Appli
698	4	13.8	82	1	US-08-042-747A-5	Sequence 5, Appli	771	4	13.8	103	1	US-08-486-013-22	Sequence 22, Appli
699	4	13.8	84	1	US-07-707-114-1	Sequence 1, Appli	772	4	13.8	103	2	US-08-482-379-20	Sequence 20, Appli
700	4	13.8	84	1	US-07-863-014-2	Sequence 2, Appli	773	4	13.8	103	2	US-08-482-279-22	Sequence 22, Appli
701	4	13.8	84	1	US-08-332-453-2	Sequence 2, Appli	774	4	13.8	103	2	US-08-342-268-20	Sequence 20, Appli
702	4	13.8	84	1	US-08-689-190-2	Sequence 2, Appli	775	4	13.8	103	2	US-08-342-268-22	Sequence 22, Appli
703	4	13.8	84	2	US-08-142-551B-1	Sequence 1, Appli	776	4	13.8	103	3	US-09-015-968-20	Sequence 20, Appli
704	4	13.8	84	2	US-08-835-231-9	Sequence 9, Appli	777	4	13.8	103	3	US-09-015-968-22	Sequence 22, Appli
705	4	13.8	84	2	US-08-411-726-1	Sequence 1, Appli	778	4	13.8	103	4	US-08-905-223-437	Sequence 437, App
706	4	13.8	84	2	US-08-805-918-3	Sequence 3, Appli	779	4	13.8	103	4	US-09-397-386-20	Sequence 20, Appli
707	4	13.8	84	3	US-09-044-536A-36	Sequence 36, Appli	780	4	13.8	103	4	US-09-397-386-22	Sequence 22, Appli
708	4	13.8	84	4	US-09-108-661-9	Sequence 9, Appli	781	4	13.8	104	4	US-08-818-112-145	Sequence 145, App
709	4	13.8	84	5	PCT-US95-15800-25	Sequence 25, Appli	782	4	13.8	104	4	US-08-818-111-140	Sequence 140, App
710	4	13.8	86	4	US-09-448-806C-3	Sequence 3, Appli	783	4	13.8	104	4	US-09-056-556-145	Sequence 145, App
711	4	13.8	87	3	US-08-906-769-161	Sequence 161, App	784	4	13.8	105	1	US-08-306-871-19	Sequence 19, Appli
712	4	13.8	87	3	US-08-906-616-161	Sequence 161, App	785	4	13.8	105	1	US-08-417-460-6	Sequence 6, Appli
713	4	13.8	87	3	US-08-639-075A-161	Sequence 161, App	786	4	13.8	105	1	US-08-569-959-19	Sequence 19, Appli
714	4	13.8	87	4	US-09-012-431-161	Sequence 161, App	787	4	13.8	105	2	US-08-553-541B-8	Sequence 8, Appli
715	4	13.8	87	4	US-09-012-692-161	Sequence 161, App	788	4	13.8	105	4	US-09-268-202-8	Sequence 8, Appli
716	4	13.8	87	4	US-08-906-613-161	Sequence 161, App	789	4	13.8	106	4	US-08-936-165A-293	Sequence 293, App
717	4	13.8	87	4	US-09-220-528-115	Sequence 115, App	790	4	13.8	107	4	US-09-220-528-52	Sequence 52, Appli
718	4	13.8	88	3	US-09-045-632-7	Sequence 7, Appli	791	4	13.8	107	4	US-08-638-610-4	Sequence 4, Appli
719	4	13.8	88	4	US-09-330-970-6	Sequence 6, Appli	792	4	13.8	109	2	US-08-429-965-4	Sequence 4, Appli
720	4	13.8	89	3	US-08-906-769-165	Sequence 165, App	793	4	13.8	109	2	US-08-969-584-4	Sequence 4, Appli
721	4	13.8	89	3	US-08-906-616-165	Sequence 165, App	794	4	13.8	109	2	US-08-970-179A-4	Sequence 4, Appli
722	4	13.8	89	4	US-08-639-075A-165	Sequence 165, App	795	4	13.8	109	2	US-08-969-249A-4	Sequence 4, Appli
723	4	13.8	89	4	US-09-012-431-165	Sequence 165, App	796	4	13.8	109	3	US-08-969-378-4	Sequence 4, Appli
724	4	13.8	89	4	US-09-012-692-165	Sequence 165, App	797	4	13.8	109	4	US-09-124-671-29	Sequence 29, Appli
725	4	13.8	89	4	US-08-906-613-165	Sequence 165, App	798	4	13.8	109	4	US-09-136-594-4	Sequence 4, Appli
726	4	13.8	93	1	US-08-466-033-69	Sequence 69, Appli	799	4	13.8	109	4	US-09-137-134-4	Sequence 4, Appli
727	4	13.8	93	2	US-08-444-733-69	Sequence 69, Appli	800	4	13.8	109	4	US-09-091-725-31	Sequence 31, Appli
728	4	13.8	93	2	US-08-464-134-69	Sequence 69, Appli	801	4	13.8	110	2	US-08-637-759B-456	Sequence 456, App
729	4	13.8	93	2	US-08-461-361-69	Sequence 69, Appli	802	4	13.8	110	3	US-08-871-355A-456	Sequence 456, App
730	4	13.8	93	2	US-08-341-843B-8	Sequence 8, Appli	803	4	13.8	110	4	US-09-201-945-456	Sequence 456, App
731	4	13.8	93	2	US-08-341-843B-20	Sequence 20, Appli	804	4	13.8	111	1	US-08-543-238-8	Sequence 8, Appli
732	4	13.8	93	2	US-08-485-910-69	Sequence 69, Appli	805	4	13.8	111	1	US-08-420-526-8	Sequence 8, Appli
733	4	13.8	93	2	US-08-427-497B-13	Sequence 13, Appli	806	4	13.8	111	1	US-08-334-773A-6	Sequence 6, Appli
734	4	13.8	93	2	US-08-427-497B-25	Sequence 25, Appli	807	4	13.8	111	1	US-08-466-886-35	Sequence 35, Appli
735	4	13.8	93	3	US-08-946-329A-77	Sequence 77, Appli	808	4	13.8	111	4	US-08-469-617-35	Sequence 35, Appli
736	4	13.8	93	4	US-09-025-596-92	Sequence 92, Appli	809	4	13.8	113	2	US-08-563-892A-26	Sequence 26, Appli
737	4	13.8	93	5	PCT-US95-06266-53	Sequence 53, Appli	810	4	13.8	113	2	US-08-563-892A-30	Sequence 30, Appli
738	4	13.8	94	1	US-08-466-033-67	Sequence 67, Appli	811	4	13.8	113	2	US-08-563-892A-31	Sequence 31, Appli
739	4	13.8	94	2	US-08-444-733-67	Sequence 67, Appli	812	4	13.8	113	2	US-08-690-011A-21	Sequence 21, Appli
740	4	13.8	94	2	US-08-464-134-67	Sequence 67, Appli	813	4	13.8	116	4	US-09-188-930-121	Sequence 121, App
741	4	13.8	94	2	US-08-461-361-67	Sequence 67, Appli	814	4	13.8	116	4	US-09-018-635-25	Sequence 25, Appli
742	4	13.8	94	2	US-08-485-910-67	Sequence 67, Appli	815	4	13.8	117	2	US-08-910-856-4	Sequence 4, Appli
743	4	13.8	94	4	US-09-486-580A-3	Sequence 3, Appli	816	4	13.8	118	3	US-09-045-764A-5	Sequence 5, Appli
744	4	13.8	94	5	PCT-US95-06266-51	Sequence 51, Appli	817	4	13.8	119	1	US-08-256-077-4	Sequence 4, Appli
745	4	13.8	97	2	US-08-726-306A-128	Sequence 128, App	818	4	13.8	119	1	US-08-466-127-4	Sequence 4, Appli
746	4	13.8	98	1	US-08-036-210-25	Sequence 25, Appli	819	4	13.8	120	4	US-09-036-315-9	Sequence 9, Appli
747	4	13.8	98	2	US-08-918-727-3	Sequence 3, Appli	820	4	13.8	121	1	US-08-463-262A-5	Sequence 5, Appli
748	4	13.8	98	2	US-08-449-609-25	Sequence 25, Appli	821	4	13.8	121	1	US-08-463-989-5	Sequence 5, Appli
749	4	13.8	98	3	US-09-205-680A-3	Sequence 3, Appli	822	4	13.8	121	1	US-08-466-033-51	Sequence 51, Appli
750	4	13.8	98	3	US-09-048-889-11	Sequence 11, Appli	823	4	13.8	121	2	US-08-444-733-51	Sequence 51, Appli
751	4	13.8	99	2	US-08-484-905-61	Sequence 61, Appli	824	4	13.8	121	2	US-08-464-134-51	Sequence 51, Appli
752	4	13.8	99	2	US-09-047-125-29	Sequence 29, Appli	825	4	13.8	121	2	US-08-461-361-51	Sequence 51, Appli
753	4	13.8	99	3	US-08-481-985B-61	Sequence 61, Appli	826	4	13.8	121	2	US-08-485-910-51	Sequence 51, Appli
754	4	13.8	99	3	US-07-736-335B-29	Sequence 29, Appli	827	4	13.8	121	2	US-08-540-804-8	Sequence 8, Appli
755	4	13.8	99	4	US-08-370-476-61	Sequence 61, Appli	828	4	13.8	121	3	US-08-218-265-8	Sequence 8, Appli
756	4	13.8	100	1	US-08-158-189-7	Sequence 7, Appli	829	4	13.8	121	3	US-08-521-872-8	Sequence 8, Appli
757	4	13.8	100	4	US-09-025-596-80	Sequence 80, Appli	830	4	13.8	121	4	US-08-590-399-8	Sequence 8, Appli

831	4	13.8	121	4	US-09-003-574-5	Sequence 5, Appli	904	4	13.8	147	2	US-08-824-405-6	Sequence 6, Appli
832	4	13.8	121	4	US-09-003-570-5	Sequence 5, Appli	905	4	13.8	147	2	US-08-824-405-12	Sequence 12, Appl
833	4	13.8	121	5	PCT-US95-06266-35	Sequence 35, Appl	906	4	13.8	147	4	US-09-347-833-10	Sequence 10, Appl
834	4	13.8	121	5	PCT-US95-13658-4	Sequence 4, Appli	907	4	13.8	148	1	US-08-256-077-2	Sequence 2, Appli
835	4	13.8	122	1	US-08-782-177-2	Sequence 2, Appli	908	4	13.8	148	1	US-08-466-127-2	Sequence 47, Appl
836	4	13.8	122	4	US-09-227-357-306	Sequence 306, App	909	4	13.8	149	1	US-07-912-740D-47	Sequence 2, Appli
837	4	13.8	123	2	US-08-626-685A-10	Sequence 10, Appl	910	4	13.8	149	1	US-08-100-874-2	Sequence 4, Appli
838	4	13.8	123	3	US-08-938-548B-10	Sequence 10, Appl	911	4	13.8	149	3	US-08-963-409-4	Sequence 4, Appli
839	4	13.8	123	4	US-08-939-093A-10	Sequence 10, Appl	912	4	13.8	149	3	US-08-641-873-20	Sequence 20, Appl
840	4	13.8	124	1	US-08-216-846-1	Sequence 1, Appli	913	4	13.8	149	4	US-09-221-357-196	Sequence 196, App
841	4	13.8	124	1	US-08-782-177-1	Sequence 1, Appli	914	4	13.8	150	2	US-08-474-379C-22	Sequence 22, Appl
842	4	13.8	124	4	US-09-199-637A-53	Sequence 53, Appl	915	4	13.8	150	3	US-09-146-249A-22	Sequence 22, Appl
843	4	13.8	125	2	US-08-862-480B-1	Sequence 1, Appli	916	4	13.8	150	3	US-08-206-188B-22	Sequence 22, Appl
844	4	13.8	126	1	US-08-312-387B-2	Sequence 2, Appli	917	4	13.8	151	2	US-08-824-405-7	Sequence 7, Appli
845	4	13.8	126	1	US-08-683-426-2	Sequence 2, Appli	918	4	13.8	151	3	US-08-916-043-5	Sequence 5, Appli
846	4	13.8	126	1	US-08-683-458-2	Sequence 2, Appli	919	4	13.8	151	6	5212286-2	Patent No. 5212286
847	4	13.8	126	2	US-08-878-360-2	Sequence 2, Appli	920	4	13.8	152	1	US-08-450-945-71	Sequence 71, Appl
848	4	13.8	126	3	US-08-478-140B-2	Sequence 2, Appli	921	4	13.8	152	3	US-08-963-409-5	Sequence 5, Appli
849	4	13.8	126	4	US-09-333-412-2	Sequence 2, Appli	922	4	13.8	152	4	US-08-976-161-71	Sequence 71, Appl
850	4	13.8	128	2	US-08-154-292B-28	Sequence 28, Appl	923	4	13.8	153	4	US-08-863-813A-34	Sequence 34, Appl
851	4	13.8	128	3	US-08-759-581B-21	Sequence 21, Appl	924	4	13.8	153	4	US-08-676-318A-34	Sequence 34, Appl
852	4	13.8	128	3	US-08-845-623-28	Sequence 28, Appl	925	4	13.8	154	1	US-08-102-942A-5	Sequence 5, Appli
853	4	13.8	128	3	US-08-815-927-28	Sequence 28, Appl	926	4	13.8	154	3	US-09-191-647-10	Sequence 10, Appl
854	4	13.8	128	4	US-08-845-258-41	Sequence 41, Appl	927	4	13.8	154	4	US-09-540-245A-10	Sequence 10, Appl
855	4	13.8	128	4	US-08-990-571-41	Sequence 41, Appl	928	4	13.8	154	4	US-09-540-153-10	Sequence 10, Appl
856	4	13.8	128	4	US-08-723-142A-41	Sequence 41, Appl	929	4	13.8	154	4	US-09-037-179B-5	Sequence 5, Appli
857	4	13.8	128	4	US-09-103-330-28	Sequence 28, Appl	930	4	13.8	155	3	US-08-819-177-4	Sequence 4, Appli
858	4	13.8	128	4	US-09-304-711-21	Sequence 21, Appl	931	4	13.8	155	4	US-08-432-994A-8	Sequence 8, Appli
859	4	13.8	129	4	US-09-199-637A-97	Sequence 97, Appl	932	4	13.8	155	4	US-09-417-455-5	Sequence 5, Appli
860	4	13.8	130	1	US-08-426-627-18	Sequence 18, Appl	933	4	13.8	155	4	US-09-348-942-5	Sequence 5, Appli
861	4	13.8	130	2	US-09-121-179-3	Sequence 3, Appli	934	4	13.8	155	4	US-09-316-081-5	Sequence 5, Appli
862	4	13.8	130	3	US-08-938-548B-6	Sequence 6, Appli	935	4	13.8	155	6	5494663-8	Patent No. 5494663
863	4	13.8	130	4	US-08-939-093A-6	Sequence 6, Appli	936	4	13.8	156	3	US-08-600-982-30	Sequence 30, Appl
864	4	13.8	131	2	US-08-385-590A-2	Sequence 2, Appli	937	4	13.8	156	5	PCT-US94-10261A-30	Sequence 30, Appl
865	4	13.8	131	2	US-08-535-864A-6	Sequence 6, Appli	938	4	13.8	159	1	US-08-253-155A-27	Sequence 27, Appl
866	4	13.8	131	3	US-09-021-520-2	Sequence 2, Appli	939	4	13.8	159	2	US-08-928-419-6	Sequence 6, Appli
867	4	13.8	132	4	US-09-224-505-5	Sequence 5, Appli	940	4	13.8	159	4	US-09-291-283-6	Sequence 6, Appli
868	4	13.8	132	4	US-09-224-505-6	Sequence 6, Appli	941	4	13.8	160	2	US-08-408-095-27	Sequence 27, Appl
869	4	13.8	133	4	US-08-468-560C-8	Sequence 8, Appli	942	4	13.8	160	3	US-09-221-235-11	Sequence 11, Appl
870	4	13.8	134	1	US-08-463-262A-8	Sequence 8, Appli	943	4	13.8	160	3	US-09-221-928-11	Sequence 11, Appl
871	4	13.8	134	1	US-08-463-989-8	Sequence 8, Appli	944	4	13.8	160	4	US-09-221-527-11	Sequence 11, Appl
872	4	13.8	134	2	US-08-850-910A-45	Sequence 45, Appl	945	4	13.8	160	4	US-09-221-236-11	Sequence 11, Appl
873	4	13.8	134	2	US-08-850-910A-48	Sequence 48, Appl	946	4	13.8	160	4	US-09-221-416-11	Sequence 11, Appl
874	4	13.8	134	2	US-09-121-179-4	Sequence 4, Appli	947	4	13.8	160	4	US-09-221-245-11	Sequence 11, Appl
875	4	13.8	134	4	US-09-273-163-8	Sequence 8, Appli	948	4	13.8	160	4	US-09-163-115-11	Sequence 11, Appl
876	4	13.8	134	4	US-09-003-574-8	Sequence 8, Appli	949	4	13.8	160	4	US-09-221-528-11	Sequence 11, Appl
877	4	13.8	134	4	US-09-003-570-8	Sequence 8, Appli	950	4	13.8	160	4	US-09-052-877-5	Sequence 5, Appli
878	4	13.8	134	5	PCT-US93-07213-15	Sequence 15, Appl	951	4	13.8	160	4	US-09-593-553-11	Sequence 11, Appl
879	4	13.8	135	1	US-08-188-281B-6	Sequence 6, Appli	952	4	13.8	160	4	US-09-117-257-38	Sequence 38, Appl
880	4	13.8	135	5	PCT-US94-07280-6	Sequence 6, Appli	953	4	13.8	160	4	US-09-221-237-11	Sequence 11, Appl
881	4	13.8	135	5	PCT-US95-01087-6	Sequence 6, Appli	954	4	13.8	160	4	US-09-489-352-38	Sequence 38, Appl
882	4	13.8	136	4	US-08-936-165A-461	Sequence 461, App	955	4	13.8	161	2	US-08-997-080-158	Sequence 158, App
883	4	13.8	136	6	5212286-4	Patent No. 5212286	956	4	13.8	161	2	US-08-997-362-158	Sequence 158, App
884	4	13.8	138	1	US-07-827-506B-2	Sequence 2, Appli	957	4	13.8	161	4	US-09-095-855-158	Sequence 158, App
885	4	13.8	138	1	US-07-827-506B-4	Sequence 4, Appli	958	4	13.8	161	4	US-09-461-697-423	Sequence 423, App
886	4	13.8	138	4	US-09-651-656-33	Sequence 33, Appl	959	4	13.8	161	4	US-09-324-542-158	Sequence 158, App
887	4	13.8	138	4	US-09-651-656-39	Sequence 39, Appl	960	4	13.8	162	1	US-08-316-239B-3	Sequence 3, Appli
888	4	13.8	139	3	US-08-753-007A-33	Sequence 33, Appl	961	4	13.8	162	1	US-08-316-239B-4	Sequence 4, Appli
889	4	13.8	139	4	US-09-398-496-33	Sequence 33, Appl	962	4	13.8	163	4	US-08-857-076-23	Sequence 23, Appl
890	4	13.8	142	1	US-07-951-715A-24	Sequence 24, Appl	963	4	13.8	164	2	US-08-824-405-2	Sequence 2, Appli
891	4	13.8	142	2	US-08-459-448A-24	Sequence 24, Appl	964	4	13.8	165	3	US-08-966-317-1	Sequence 1, Appli
892	4	13.8	142	3	US-08-459-595A-24	Sequence 24, Appl	965	4	13.8	166	5	PCT-US94-04361-44	Sequence 44, Appl
893	4	13.8	142	3	US-08-459-504B-24	Sequence 24, Appl	966	4	13.8	168	2	US-08-612-788-27	Sequence 27, Appl
894	4	13.8	142	3	US-08-459-444-24	Sequence 24, Appl	967	4	13.8	168	3	US-09-066-028-27	Sequence 27, Appl
895	4	13.8	142	4	US-09-547-422-24	Sequence 24, Appl	968	4	13.8	168	5	PCT-US94-04361-43	Sequence 43, Appl
896	4	13.8	144	2	US-08-824-405-13	Sequence 13, Appl	969	4	13.8	170	2	US-08-652-558-40	Sequence 40, Appl
897	4	13.8	144	5	PCT-US93-07213-11	Sequence 11, Appl	970	4	13.8	170	4	US-09-002-114-1	Sequence 1, Appli
898	4	13.8	144	6	5202239-1	Patent No. 5202239	971	4	13.8	171	1	US-08-313-681A-4	Sequence 4, Appli
899	4	13.8	144	6	5202239-3	Patent No. 5202239	972	4	13.8	171	3	US-09-322-911-4	Sequence 4, Appli
900	4	13.8	145	3	US-09-100-391-8	Sequence 8, Appli	973	4	13.8	171	4	US-08-611-587-5	Sequence 5, Appli
901	4	13.8	145	4	US-08-720-625-5	Sequence 5, Appli	974	4	13.8	172	1	US-08-188-281B-5	Sequence 5, Appli
902	4	13.8	145	5	PCT-US94-01149-69	Sequence 69, Appl	975	4	13.8	172	1	US-08-438-753B-2	Sequence 2, Appli
903	4	13.8	145	5	PCT-US94-01149-70	Sequence 70, Appl	976	4	13.8	172	1	US-08-443-883A-2	Sequence 2, Appli

977	4	13.8	172	2	US-08-631-328-2	Sequence 2, Appli	1050	4	13.8	188	2	US-09-004-713-3	Sequence 3, Appli
978	4	13.8	172	2	US-08-455-524B-2	Sequence 2, Appli	1051	4	13.8	189	1	US-08-152-485-4	Sequence 4, Appli
979	4	13.8	172	2	US-08-455-021B-2	Sequence 2, Appli	1052	4	13.8	189	1	US-08-463-089-4	Sequence 4, Appli
980	4	13.8	172	3	US-08-860-165-12	Sequence 12, Appl	1053	4	13.8	189	1	US-08-461-360A-4	Sequence 4, Appli
981	4	13.8	172	4	US-09-045-467-2	Sequence 2, Appli	1054	4	13.8	189	1	US-08-461-359-4	Sequence 4, Appli
982	4	13.8	172	4	US-08-954-395A-18	Sequence 18, Appl	1055	4	13.8	189	5	PCT-US94-12904-4	Sequence 4, Appli
983	4	13.8	172	4	US-09-359-382-12	Sequence 15, Appl	1056	4	13.8	190	2	US-08-531-525-45	Sequence 4, Appli
984	4	13.8	172	5	PCT-US94-07280-5	Sequence 5, Appli	1057	4	13.8	190	2	US-08-845-998-4	Sequence 4, Appli
985	4	13.8	172	5	PCT-US95-01087-5	Sequence 5, Appli	1058	4	13.8	190	2	US-08-845-998-6	Sequence 6, Appli
986	4	13.8	173	2	US-08-467-046-4	Sequence 4, Appli	1059	4	13.8	190	2	US-08-718-270A-45	Sequence 4, Appli
987	4	13.8	173	3	US-08-669-408B-8	Sequence 8, Appli	1060	4	13.8	190	4	US-09-206-537-4	Sequence 4, Appli
988	4	13.8	173	4	US-08-858-207A-365	Sequence 365, App	1061	4	13.8	190	4	US-09-206-537-6	Sequence 6, Appli
989	4	13.8	174	1	US-08-428-732-2	Sequence 2, Appli	1062	4	13.8	190	4	US-09-430-854-4	Sequence 4, Appli
990	4	13.8	174	3	US-09-122-443-6	Sequence 6, Appli	1063	4	13.8	190	4	US-09-430-854-6	Sequence 6, Appli
991	4	13.8	174	3	US-09-122-443-7	Sequence 7, Appli	1064	4	13.8	191	3	US-08-928-361B-26	Sequence 26, Appl
992	4	13.8	174	4	US-09-383-586-12	Sequence 12, Appl	1065	4	13.8	192	1	US-08-086-428B-52	Sequence 52, Appl
993	4	13.8	174	4	US-09-199-637A-63	Sequence 63, Appl	1066	4	13.8	192	1	US-08-086-428B-54	Sequence 54, Appl
994	4	13.8	175	1	US-08-428-732-5	Sequence 5, Appli	1067	4	13.8	192	1	US-08-086-428B-55	Sequence 55, Appl
995	4	13.8	175	1	US-09-230-637-24	Sequence 24, Appl	1068	4	13.8	192	1	US-08-086-428B-59	Sequence 59, Appl
996	4	13.8	176	1	US-08-415-751-1	Sequence 1, Appli	1069	4	13.8	192	1	US-08-440-103-40	Sequence 40, Appl
997	4	13.8	176	4	US-09-232-446B-4	Sequence 4, Appli	1070	4	13.8	192	1	US-08-440-103-45	Sequence 45, Appl
998	4	13.8	176	4	US-09-213-343-3	Sequence 3, Appli	1071	4	13.8	192	1	US-08-440-542-40	Sequence 40, Appl
999	4	13.8	177	1	US-08-181-271A-46	Sequence 46, Appl	1072	4	13.8	192	1	US-08-440-542-45	Sequence 45, Appl
1000	4	13.8	177	1	US-08-449-315-46	Sequence 46, Appl	1073	4	13.8	192	1	US-08-231-368-40	Sequence 40, Appl
1001	4	13.8	177	1	US-08-444-803-46	Sequence 46, Appl	1074	4	13.8	192	1	US-08-231-368-45	Sequence 45, Appl
1002	4	13.8	177	1	US-08-449-043-46	Sequence 46, Appl	1075	4	13.8	192	1	US-08-440-210-40	Sequence 40, Appl
1003	4	13.8	177	1	US-08-456-265A-46	Sequence 46, Appl	1076	4	13.8	192	1	US-08-440-210-45	Sequence 45, Appl
1004	4	13.8	177	1	US-08-456-736-46	Sequence 46, Appl	1077	4	13.8	192	2	US-08-468-570-52	Sequence 52, Appl
1005	4	13.8	177	1	US-08-455-416-46	Sequence 46, Appl	1078	4	13.8	192	2	US-08-468-570-54	Sequence 54, Appl
1006	4	13.8	177	1	US-08-455-244-46	Sequence 46, Appl	1079	4	13.8	192	2	US-08-468-570-55	Sequence 55, Appl
1007	4	13.8	177	1	US-08-454-876-46	Sequence 46, Appl	1080	4	13.8	192	2	US-08-468-570-59	Sequence 59, Appl
1008	4	13.8	177	2	US-08-457-364-46	Sequence 46, Appl	1081	4	13.8	192	2	US-08-290-665A-52	Sequence 52, Appl
1009	4	13.8	177	2	US-08-456-262-46	Sequence 46, Appl	1082	4	13.8	192	2	US-08-290-665A-54	Sequence 54, Appl
1010	4	13.8	177	2	US-08-456-240-46	Sequence 46, Appl	1083	4	13.8	192	2	US-08-290-665A-55	Sequence 55, Appl
1011	4	13.8	177	2	US-08-455-736-46	Sequence 46, Appl	1084	4	13.8	192	2	US-08-290-665A-59	Sequence 59, Appl
1012	4	13.8	177	2	US-08-971-217-46	Sequence 46, Appl	1085	4	13.8	192	4	US-08-444-818-165	Sequence 165, App
1013	4	13.8	177	2	US-08-921-382-2	Sequence 2, Appli	1086	4	13.8	192	4	US-08-444-818-166	Sequence 166, App
1014	4	13.8	177	2	US-08-967-364-5	Sequence 5, Appli	1087	4	13.8	192	4	US-09-046-604-40	Sequence 40, Appl
1015	4	13.8	177	3	US-08-967-364-9	Sequence 9, Appli	1088	4	13.8	192	4	US-09-046-604-45	Sequence 45, Appl
1016	4	13.8	177	3	US-09-368-408-5	Sequence 9, Appli	1089	4	13.8	192	4	PCT-US95-10398-52	Sequence 52, Appl
1017	4	13.8	177	3	US-09-368-408-9	Sequence 9, Appli	1090	4	13.8	192	5	PCT-US95-10398-54	Sequence 54, Appl
1018	4	13.8	178	4	US-09-350-600-46	Sequence 46, Appl	1091	4	13.8	192	5	PCT-US95-10398-55	Sequence 55, Appl
1019	4	13.8	179	3	US-09-171-461-2	Sequence 2, Appli	1092	4	13.8	192	5	PCT-US95-10398-59	Sequence 59, Appl
1020	4	13.8	180	1	US-09-042-771-4	Sequence 4, Appli	1093	4	13.8	192	5	PCT-US95-10398-59	Sequence 59, Appl
1021	4	13.8	180	1	US-07-949-812-5	Sequence 5, Appli	1094	4	13.8	193	2	US-08-861-549-4	Sequence 4, Appli
1022	4	13.8	181	3	US-08-447-010-7	Sequence 7, Appli	1095	4	13.8	193	2	US-08-896-933-31	Sequence 31, Appl
1023	4	13.8	181	3	US-08-753-007A-4	Sequence 4, Appli	1096	4	13.8	193	4	US-09-314-235-31	Sequence 31, Appl
1024	4	13.8	181	4	US-09-398-496-4	Sequence 4, Appli	1097	4	13.8	194	1	US-08-446-918A-6	Sequence 6, Appli
1025	4	13.8	181	4	US-09-320-528-40	Sequence 40, Appl	1098	4	13.8	194	2	US-08-435-605A-6	Sequence 6, Appli
1026	4	13.8	181	4	US-08-858-207A-317	Sequence 317, App	1099	4	13.8	194	2	US-08-435-605A-14	Sequence 14, Appl
1027	4	13.8	182	4	US-09-199-637A-327	Sequence 327, App	1100	4	13.8	194	2	US-08-435-605A-16	Sequence 16, Appl
1028	4	13.8	183	2	US-08-980-329C-2	Sequence 2, Appli	1101	4	13.8	194	2	US-08-435-605A-54	Sequence 54, Appl
1029	4	13.8	183	4	US-09-117-257-42	Sequence 42, Appl	1102	4	13.8	194	2	US-08-435-605A-55	Sequence 55, Appl
1030	4	13.8	183	4	US-09-489-352-42	Sequence 42, Appl	1103	4	13.8	194	2	US-08-580-806-6	Sequence 6, Appli
1031	4	13.8	185	1	US-08-233-389C-1	Sequence 1, Appli	1104	4	13.8	194	4	US-09-117-257-17	Sequence 17, Appl
1032	4	13.8	185	2	US-08-801-863-1	Sequence 1, Appli	1105	4	13.8	194	4	US-09-489-352-17	Sequence 17, Appl
1033	4	13.8	185	2	US-08-486-596A-1	Sequence 1, Appli	1106	4	13.8	195	4	US-09-029-213B-23	Sequence 23, Appl
1034	4	13.8	185	3	US-09-004-713-1	Sequence 1, Appli	1107	4	13.8	195	4	US-09-117-257-34	Sequence 34, Appl
1035	4	13.8	185	3	US-09-045-632-26	Sequence 26, Appl	1108	4	13.8	195	4	US-09-489-352-34	Sequence 34, Appl
1036	4	13.8	185	4	US-09-154-083-14	Sequence 14, Appl	1109	4	13.8	195	4	US-09-117-257-34	Sequence 34, Appl
1037	4	13.8	185	4	US-09-056-556-232	Sequence 232, App	1110	4	13.8	198	4	US-09-228-986-93	Sequence 93, Appl
1038	4	13.8	185	6	5489533-4	Patent No. 5489533	1111	4	13.8	199	2	US-08-665-647-9	Sequence 9, Appli
1039	4	13.8	186	4	US-08-858-207A-330	Sequence 330, App	1112	4	13.8	199	2	US-08-474-379C-38	Sequence 38, Appl
1040	4	13.8	186	4	US-09-199-637A-7	Sequence 7, Appli	1113	4	13.8	199	3	US-09-146-249A-38	Sequence 38, Appl
1041	4	13.8	187	1	US-08-049-473-33	Sequence 33, Appl	1114	4	13.8	199	3	US-08-206-188B-38	Sequence 38, Appl
1042	4	13.8	187	1	US-08-312-648-33	Sequence 33, Appl	1115	4	13.8	199	4	US-08-995-159-2	Sequence 2, Appli
1043	4	13.8	187	1	US-07-956-700B-17	Sequence 17, Appl	1116	4	13.8	199	4	US-08-995-159-25	Sequence 25, Appl
1044	4	13.8	187	1	US-08-476-537-17	Sequence 17, Appl	1117	4	13.8	200	2	US-08-477-396A-19	Sequence 19, Appl
1045	4	13.8	187	2	US-08-485-607-17	Sequence 17, Appl	1118	4	13.8	200	3	US-09-248-335-72	Sequence 72, Appl
1046	4	13.8	187	2	US-08-475-879-17	Sequence 17, Appl	1119	4	13.8	200	5	PCT-US93-07213-7	Sequence 7, Appli
1047	4	13.8	188	5	PCT-US94-04190-33	Sequence 33, Appl	1120	4	13.8	201	1	US-07-688-352C-38	Sequence 38, Appl
1048	4	13.8	188	2	US-08-233-389C-3	Sequence 3, Appli	1121	4	13.8	201	5	PCT-US91-02714-37	Sequence 37, Appl
1049	4	13.8	188	2	US-08-801-863-3	Sequence 3, Appli	1122	4	13.8	202	2	US-08-948-616-11	Sequence 11, Appl


```
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF
; FILE REFERENCE: 5820.546
; CURRENT APPLICATION NUMBER: US/09/374,493
; CURRENT FILING DATE: 1999-08-13
; EARLIER APPLICATION NUMBER: 09/150,133
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 60/072,994
; EARLIER FILING DATE: 1998-01-29
; EARLIER APPLICATION NUMBER: PCT/US99/16750
; EARLIER FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 8.0 (saved in ASCII format)
; SEQ ID NO 11
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-374-493-11

Query Match          20.7%; Score 6; DB 4; Length 359;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SELAPR 20
Db 161 SELAPR 166

RESULT 4
US-09-374-824-11
; Sequence 11, Application US/09374824
; Patent No. 6207414
; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF
; FILE REFERENCE: 5820.547
; CURRENT APPLICATION NUMBER: US/09/374,824
; CURRENT FILING DATE: 1999-08-13
; EARLIER APPLICATION NUMBER: 09/150,133
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 60/072,994
; EARLIER FILING DATE: 1998-01-29
; EARLIER APPLICATION NUMBER: PCT/US99/16750
; EARLIER FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 8.0 (saved in ASCII format)
; SEQ ID NO 11
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-374-824-11

Query Match          20.7%; Score 6; DB 4; Length 359;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SELAPR 20
Db 161 SELAPR 166

RESULT 5
US-09-374-492-11
; Sequence 11, Application US/09374492
; Patent No. 6207432
; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: TYROSYLPROTEIN SULFOTRANSFERASES AND METHODS OF USE THEREOF
; FILE REFERENCE: 5820.545
; CURRENT APPLICATION NUMBER: US/09/374,492
; CURRENT FILING DATE: 1999-08-13
```

```
; EARLIER APPLICATION NUMBER: 09/150,141
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 60/072,994
; EARLIER FILING DATE: 1998-01-29
; EARLIER APPLICATION NUMBER: PCT/US99/16750
; EARLIER FILING DATE: 1999-07-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Wordperfect 8.0 (saved in ASCII format)
; SEQ ID NO 11
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-374-492-11

Query Match          20.7%; Score 6; DB 4; Length 359;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SELAPR 20
Db 161 SELAPR 166

RESULT 6
US-08-720-899-7
; Sequence 7, Application US/08720899
; Patent No. 5753460
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 5753460o No. 5753460disk of No. 5753460th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,899
; FILING DATE: 10-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Jr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-720-899-7

Query Match          20.7%; Score 6; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 38;
```


Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSSEL 17
Db 353 PTDSSEL 358

RESULT 7

US-08-459-610-7
; Sequence 7, Application US/08459610
; Patent No. 5801043
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thelliersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5801043o No. 5801043disk of No. 5801043th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,610
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr. Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-459-610-7

Query Match 20.7%; Score 6; DB 1; Length 478;

Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSSEL 17
Db 353 PTDSSEL 358

RESULT 8

US-08-343-804-7
; Sequence 7, Application US/08343804
; Patent No. 5830837
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thelliersen, Marianne
; APPLICANT: Van der Zee, Pia

; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5830837o No. 5830837disk of No. 5830837th America, Inc.

; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/343,804
; FILING DATE: 22-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr. Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-343-804-7

Query Match 20.7%; Score 6; DB 2; Length 478;

Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSSEL 17
Db 353 PTDSSEL 358

RESULT 9

US-08-339-715A-2
; Sequence 2, Application US/08339715A
; Patent No. 5965442
; GENERAL INFORMATION:
; APPLICANT: Kaneko, Hiroki
; APPLICANT: Takada, Toshikazu
; APPLICANT: Shimada, Jiro
; APPLICANT: Kuriki, Takashi
; APPLICANT: Yanase, Michiyo
; APPLICANT: Takata, Hiroki
; APPLICANT: Okada, Shigetaka
; TITLE OF INVENTION: METHOD OF ALTERING ENZYMES AND A NOVEL
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: POPHAM, HAIK, SCHNOBRICH & KAUFMAN, LTD.
; STREET: Metropolitan Square Building, Suite 800, 1450
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/339,715A

Query Match 20.7%; Score 6; DB 1; Length 478;

Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSSEL 17
Db 353 PTDSSEL 358

RESULT 8

US-08-343-804-7
; Sequence 7, Application US/08343804
; Patent No. 5830837
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thelliersen, Marianne
; APPLICANT: Van der Zee, Pia

; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 306096/1993
; FILING DATE: 12-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Melsner, Allen S.
; REGISTRATION NUMBER: 27,215
; REFERENCE/DOCKET NUMBER: 18335.009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-824-8000
; TELEFAX: 202-824-8199
; TELEX: 248516
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-339-715A-2

Query Match 20.7%; Score 6; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 PTDSEL 17
Db 353 PTDSEL 358

RESULT 10
US-08-600-908A-10
; Sequence 10, Application US/08600908A
; Patent No. 5989169
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisg rd-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: '-Amylase Mutants
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5989169o No. 5989169disk of No. 5989169th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,908A
; FILING DATE: 13-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 4394.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-08-600-908A-10

Query Match 20.7%; Score 6; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 PTDSEL 17
Db 353 PTDSEL 358

RESULT 11
US-08-683-838A-10
; Sequence 10, Application US/08683838A
; Patent No. 6022724
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisg rd-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: '-Amylase Mutants
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6022724o No. 6022724disk of No. 6022724th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,838A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/600,908
; FILING DATE: 13-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 4394.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-08-683-838A-10

Query Match 20.7%; Score 6; DB 3; Length 478;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 PTDSEL 17
Db 353 PTDSEL 358

RESULT 12
US-09-182-859-7
; Sequence 7, Application US/09182859
; Patent No. 6143708
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Borchert, Torben

APPLICANT: Bisgard-Prantzen, Henrik
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 4796.204-US
; CURRENT APPLICATION NUMBER: US/09/182,859
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: 0515/96
; EARLIER FILING DATE: 1996-04-30
; EARLIER APPLICATION NUMBER: 0712/96
; EARLIER FILING DATE: 1996-06-28
; EARLIER APPLICATION NUMBER: 0775/96
; EARLIER FILING DATE: 1996-07-11
; EARLIER APPLICATION NUMBER: 1263/96
; EARLIER FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-09-182-859-7

Query Match 20.7%; Score 6; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSEL 17
Db 353 PTDSEL 358

RESULT 13
US-08-453-695A-115
; Sequence 115, Application US/08453695A
; Patent No. 5708143
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: Protocadherin Materials and Methods
; NUMBER OF SEQUENCES: 115
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
; ADDRESS: Borun
; STREET: 233 South Wacker, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,695A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: NO. 5708143and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 32658
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 115:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 616 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-453-695A-115

Query Match 20.7%; Score 6; DB 1; Length 616;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELAPRS 21
Db 372 ELAPRS 377

RESULT 14
US-08-268-161A-115
; Sequence 115, Application US/08268161A
; Patent No. 5798224
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: Protocadherin Materials and Methods
; NUMBER OF SEQUENCES: 115
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
; ADDRESS: Borun
; STREET: 233 South Wacker, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/268,161A
; FILING DATE: June 27, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Young J. Suh
; REGISTRATION NUMBER: P-41,337
; REFERENCE/DOCKET NUMBER: 27866/32149
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 115:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 616 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-268-161A-115

Query Match 20.7%; Score 6; DB 1; Length 616;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELAPRS 21
Db 372 ELAPRS 377

RESULT 15
US-08-453-702A-115
; Sequence 115, Application US/08453702A
; Patent No. 5891706
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: Protocadherin Materials and Methods
; NUMBER OF SEQUENCES: 115
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
; ADDRESS: Borun
; STREET: 233 South Wacker, 6300 Sears Tower
; CITY: Chicago

; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453.702A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5891706and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 32657
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 115:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 616 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-453-702A-115

Query Match 20.7%; Score 6; DB 2; Length 616;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 ELAPRS 21
Db 372 ELAPRS 377

RESULT 16
US-09-099-639-115
; Sequence 115, Application US/09099639
; Patent No. 6262237
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: Protocadherin Materials and Methods
; NUMBER OF SEQUENCES: 115
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
; STREET: 233 South Wacker, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/099,639
; FILING DATE: 18 JUN 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/263,161
; FILING DATE: 27 JUN 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Greta E. No. 6262237and
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/34703
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 115:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 616 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-099-639-115

Query Match 20.7%; Score 6; DB 4; Length 616;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 ELAPRS 21
Db 372 ELAPRS 377

RESULT 17
PCT-US95-08071-115
; Sequence 115, Application PC/TUS9508071
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: Protocadherin Materials and Methods
; NUMBER OF SEQUENCES: 115
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08071
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/12588
; FILING DATE: 23 DEC 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/998,003
; FILING DATE: 29 DEC 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Noland, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 32149
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 115:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 616 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-08071-115

Query Match 20.7%; Score 6; DB 5; Length 616;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 ELAPRS 21
Db 372 ELAPRS 377

RESULT 18

US-08-473-553A-5
; Sequence 5, Application US/08473553A
; Patent No. 5859338
; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliot M.
; APPLICANT: Clark, Steven E.
; APPLICANT: Williams, Robert W.
; TITLE OF INVENTION: Plant Clavatal Nucleic Acids,
; TITLE OF INVENTION: Transformed Plants, and Proteins
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,553A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-60886/RET/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 999 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-473-553A-5

Query Match 20.7%; Score 6; DB 2; Length 999;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSEL 17
Db 897 PTDSEL 902

RESULT 19

US-08-602-999A-395
; Sequence 395, Application US/08602999A
; Patent No. 6184205
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: QUILLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas

; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,999A
; FILING DATE: 16-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 395:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-602-999A-395

Query Match 17.2%; Score 5; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELAPR 20
Db 1 ELAPR 5

RESULT 20

US-08-768-147B-5
; Sequence 5, Application US/08768147B
; Patent No. 5859222
; GENERAL INFORMATION:
; APPLICANT: Leung, David W.
; APPLICANT: Tompkins, Christopher
; TITLE OF INVENTION: PHOSPHATIDYLCHOLINE PHOSPHOLIPASE D
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cell Therapeutics, Inc.
; STREET: 200 Elliott Avenue West, Suite 400
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98119
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" disk, 1.44Mb, double side, high density
; COMPUTER: Hewlett Packard-IBM Compatible
; OPERATING SYSTEM: MS-DOS Version 6
; SOFTWARE: WORD for WINDOWS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/768,147B
; FILING DATE: 16-Dec-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Provisional US60008768
; FILING DATE: 15-Dec-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Faciszewski, Stephen
; REGISTRATION NUMBER: 36,131,
; REFERENCE/DOCKET NUMBER: 1802A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)282-7100

TELEFAX: (206)284-6206
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 19
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: polypeptide
HYPOTHETICAL: no
ANTI-SENSE: no
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: yeast
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY: polypeptide fragment
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 5:
US-08-768-147B-5

Query Match 17.2%; Score 5; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 DSELA 18
Db 9 DSELA 13

RESULT 21
US-09-107-149-6
Sequence 6, Application US/09107149
Patent No. 6274363
GENERAL INFORMATION:
APPLICANT: Leung, David W.
APPLICANT: Tompkins, Christopher K.
TITLE OF INVENTION: PHOSPHATIDYLCHOLINE PHOSPHOLIPASE D
FILE REFERENCE: 077319/0144
CURRENT APPLICATION NUMBER: US/09/107,149
CURRENT FILING DATE: 1998-06-30
EARLIER APPLICATION NUMBER: 08/768,147
EARLIER FILING DATE: 1996-12-17
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6

LENGTH: 19
TYPE: PRT
ORGANISM: R83570/261972
US-09-107-149-6

Query Match 17.2%; Score 5; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 DSELA 18
Db 9 DSELA 13

RESULT 22
US-09-082-279B-835
Sequence 835, Application US/09082279B
Patent No. 6258782
GENERAL INFORMATION:
APPLICANT: Barney, Shawn
APPLICANT: Guthrie, Kelly
APPLICANT: Merutka, Gene
APPLICANT: Anwer, Mohamed
APPLICANT: Lambert, Dennis
TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC PROPERTIES
FILE REFERENCE: 7872-043
CURRENT APPLICATION NUMBER: US/09/082,279B
CURRENT FILING DATE: 1998-05-20
NUMBER OF SEQ ID NOS: 1515
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 835
LENGTH: 35
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Core polypeptide
US-09-082-279B-835

Query Match 17.2%; Score 5; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 INQLL 8
Db 18 INQLL 22

RESULT 23
US-09-315-304B-835
Sequence 835, Application US/09315304B
Patent No. 6348568
GENERAL INFORMATION:
APPLICANT: Barney, S.
APPLICANT: Guthrie, K.
APPLICANT: Merutka, G.
APPLICANT: Anwer, M.
APPLICANT: Lambert, D.
TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC PROPERTIES
FILE REFERENCE: 7872-052
CURRENT APPLICATION NUMBER: US/09/315,304B
CURRENT FILING DATE: 1999-05-20
PRIOR APPLICATION NUMBER: 09/082,279
PRIOR FILING DATE: 1998-05-20
NUMBER OF SEQ ID NOS: 1667
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 835
LENGTH: 35
TYPE: PRT
ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: Core polypeptide
US-09-315-304B-835

Query Match 17.2%; Score 5; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 INQLL 8
| | | | |
DB 18 INQLL 22

RESULT 24
US-08-905-223-384
Sequence 384, Application US/08905223
Patent No. 6222029
GENERAL INFORMATION:
APPLICANT: Edwards, Jean-Baptiste D.
APPLICANT: Duclert, Aymeric
APPLICANT: Lacroix, Bruno
TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
NUMBER OF SEQUENCES: 503
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 501 West Broadway
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-3505
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/905,223
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 384:
SEQUENCE CHARACTERISTICS:
LENGTH: 51 amino acids
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Brain
FEATURE:
NAME/KEY: sig peptide
LOCATION: -36...1
IDENTIFICATION METHOD: Von Heijne matrix
OTHER INFORMATION: score 3.9
OTHER INFORMATION: seq SLCLSLIPGPKP/LV
US-08-905-223-384

Query Match 17.2%; Score 5; DB 4; Length 51;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 INQLL 8
| | | | |
DB 42 INQLL 46

RESULT 25
US-09-047-125-5
Sequence 5, Application US/09047125
Patent No. 5976787
GENERAL INFORMATION:
APPLICANT: Leland F. Velicer, Peter Brunovskis,
TITLE OF INVENTION: Marek's Disease Herpesvirus
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ian C. McLeod
STREET: 2190 Commons Parkway
CITY: Okemos
STATE: Michigan
COUNTRY: USA
ZIP: 48864
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 1.2 Mb
COMPUTER: IBM PS2, Model 50
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: PC-Write 3.02
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/047,125
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/736,335
FILING DATE: July 25, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Ian C. McLeod
REGISTRATION NUMBER: 20,931
REFERENCE/DOCKET NUMBER: MSU 4.1-132
TELECOMMUNICATION INFORMATION:
TELEPHONE: (517) 347-4100
TELEFAX: (517) 347-4103
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 99 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal fragment
ORIGINAL SOURCE:
ORGANISM: herpes simplex virus type 1 (HSV1)
FEATURE:
NAME/KEY: peptide of HSV1 gene US1 polypeptide
LOCATION: 159 TO 257
OTHER INFORMATION: peptide homologous to the US1 gene
OTHER INFORMATION: polypeptide of MDV
US-09-047-125-5

Query Match 17.2%; Score 5; DB 2; Length 99;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CINQL 7
| | | | |
DB 39 CINQL 43

RESULT 26
US-07-736-335E-5
Sequence 5, Application US/07736335E
Patent No. 6087127
GENERAL INFORMATION:
APPLICANT: Leland F. Velicer, Peter Brunovskis,
TITLE OF INVENTION: Marek's Disease Herpesvirus
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ian C. McLeod
STREET: 2190 Commons Parkway

```

; CITY: Okemos
; STATE: Michigan
; COUNTRY: USA
; ZIP: 48864
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 1.2 Mb
; COMPUTER: IBM PS2, Model 50
; OPERATING SYSTEM: MS-DOS 5.0
; SOFTWARE: PC-Write 3.02
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/736,335E
; FILING DATE: July 25, 1991
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Ian C. McLeod
; REGISTRATION NUMBER: 20,931
; REFERENCE/DOCKET NUMBER: MSU 4.1-132
; TELEPHONE: (517) 347-4100
; TELEFAX: (517) 347-4103
;
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal fragment
; ORIGINAL SOURCE:
; ORGANISM: herpes simplex virus type 1 (HSV1)
; FEATURE:
; NAME/KEY: peptide of HSV1 gene US1 polypeptide
; LOCATION: 159 TO 257
; OTHER INFORMATION: peptide homologous to the US1 gene
; OTHER INFORMATION: polypeptide of MDV
;
; US-07-736-335E-5
;
; Query Match 17.2%; Score 5; DB 3; Length 99;
; Best Local Similarity 100.0%; Pred. No. 1.1e+02;
; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 3 CINOL 7
; Db 39 CINOL 43
;
; RESULT 27
; US-08-505-617-6
; Sequence 6, Application US/08505617
; Patent No. 5861378
; GENERAL INFORMATION:
; APPLICANT: IWANAGA, Sadaaki
; APPLICANT: KAWABATA, Shun-ichiro
; APPLICANT: SAITO, Tetsu
; TITLE OF INVENTION: POLYPEPTIDES, AND PREPARATION AND DNA
; TITLE OF INVENTION: ENCODING THEREOF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault, LLP
; STREET: 125 High Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/018,170
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/505,617
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, Paula A
; REGISTRATION NUMBER: 32,503
; REFERENCE/DOCKET NUMBER: FJN-041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-09-018-170-6

```

```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/505,617
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, Paula A
; REGISTRATION NUMBER: 32,503
; REFERENCE/DOCKET NUMBER: FJN-041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-505-617-6
;
; Query Match 17.2%; Score 5; DB 2; Length 117;
; Best Local Similarity 100.0%; Pred. No. 1.3e+02;
; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 11 LPTDS 15
; Db 17 LPTDS 21
;
; RESULT 28
; US-09-018-170-6
; Sequence 6, Application US/09018170
; Patent No. 5965725
; GENERAL INFORMATION:
; APPLICANT: IWANAGA, Sadaaki
; APPLICANT: KAWABATA, Shun-ichiro
; APPLICANT: SAITO, Tetsu
; TITLE OF INVENTION: POLYPEPTIDES, AND PREPARATION AND DNA
; TITLE OF INVENTION: ENCODING THEREOF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault, LLP
; STREET: 125 High Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/018,170
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/505,617
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, Paula A
; REGISTRATION NUMBER: 32,503
; REFERENCE/DOCKET NUMBER: FJN-041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-09-018-170-6

```


Query Match 17.2%; Score 5; DB 2; Length 117;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 LPTDS 15
 Db 17 LPTDS 21

RESULT 29
 US-08-266-451B-18
 ; Sequence 18, Application US/08266451B
 ; Patent No. 5623054
 ; GENERAL INFORMATION:
 ; APPLICANT: Zhang et al.
 ; TITLE OF INVENTION: CRUCIFER AFT PROTEINS AND USES
 ; TITLE OF INVENTION: THEREOF
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 ; COMPUTER: IBM PS/2 Model 502 or 55SX
 ; OPERATING SYSTEM: MS-DOS (Version 5.0)
 ; SOFTWARE: WordPerfect (Version 5.1)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/266.451B
 ; FILING DATE: 23-June-1994
 ; CLASSIFICATION: 800
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lech, Karen F.
 ; REGISTRATION NUMBER: 35,238
 ; REFERENCE/DOCKET NUMBER: 00786/219001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 542-5070
 ; TELEFAX: (617) 542-8906
 ; TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 18:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 122
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-266-451B-18

Query Match 17.2%; Score 5; DB 1; Length 122;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KLPTD 14
 Db 36 KLPTD 40

RESULT 30
 US-08-748-725-18
 ; Sequence 18, Application US/08748725
 ; Patent No. 5859346
 ; GENERAL INFORMATION:
 ; APPLICANT: Zhang et al.
 ; TITLE OF INVENTION: CRUCIFER AFT PROTEINS AND USES
 ; TITLE OF INVENTION: THEREOF
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: U.S.A.
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 COMPUTER: IBM PS/2 Model 502 or 55SX
 OPERATING SYSTEM: MS-DOS (Version 5.0)
 SOFTWARE: WordPerfect (Version 5.1)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/748,725
 FILING DATE:
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/266,451
 FILING DATE: 23-June-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Lech, Karen F.
 REGISTRATION NUMBER: 35,238
 REFERENCE/DOCKET NUMBER: 00786/219002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 122
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-748-725-18

Query Match 17.2%; Score 5; DB 2; Length 122;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KLPTD 14
 Db 36 KLPTD 40

RESULT 31
 US-09-188-930-187
 ; Sequence 187, Application US/09188930A
 ; Patent No. 6150502
 ; GENERAL INFORMATION:
 ; APPLICANT: Watson, James D.
 ; APPLICANT: Strachan, Lorna
 ; APPLICANT: Sleeman, Matthew
 ; APPLICANT: Onrust, Rene
 ; APPLICANT: Murison, James Greg
 ; TITLE OF INVENTION: Compositions Isolated From Skin Cells
 ; TITLE OF INVENTION: and Methods For Their Use
 ; FILE REFERENCE: 11000.1011c1
 ; CURRENT APPLICATION NUMBER: US/09/188,930A
 ; CURRENT FILING DATE: 1998-11-09
 ; NUMBER OF SEQ ID NOS: 348
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 187
 ; LENGTH: 152
 ; TYPE: PRT
 ; ORGANISM: mouse
 ; US-09-188-930-187

Query Match 17.2%; Score 5; DB 4; Length 152;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRCIN 5
 Db 1 KRCIN 5

Db 134 KRCIN 138

RESULT 32

US-09-199-637A-169
; Sequence 169, Application US/09199637A
; Patent No. 6355411

; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick

; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.

; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah

; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana

; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID

; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002

; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25

; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25

; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 169
; LENGTH: 159

; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa

US-09-199-637A-169

Query Match

Best Local Similarity 17.2%; Score 5; DB 4; Length 159;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELAPR 20

Db 150 ELAPR 154

RESULT 33

US-07-949-812-4

; Sequence 4, Application US/07949812

; Patent No. 5668007

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: RECOMBINANT 21KD COCOA PROTEIN AND PRECURSOR

; NUMBER OF SEQUENCES: 28

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/949,812

; FILING DATE: 27-JAN-1993

; CLASSIFICATION: 435

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 181 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; ORIGINAL SOURCE:

; ORGANISM: BARLEY

; FEATURE:

; NAME/KEY: Protein

; LOCATION: 1..181

; OTHER INFORMATION: /note= "ALPHA-AMYLASE-SUBSTITISIN INHIBITOR"

US-07-949-812-4

Query Match 17.2%; Score 5; DB 1; Length 181;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 DSELA 18

Db 99 DSELA 103

RESULT 34

US-08-441-629-4

; Sequence 4, Application US/08441629

; Patent No. 5766923

; GENERAL INFORMATION:

; APPLICANT: Kirschner, Marc W.

; APPLICANT: Kinoshita, No. 5766923iyuki

; TITLE OF INVENTION: RECEPTOR-LIGAND ASSAY

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

; STREET: Two Militia Drive

; CITY: Lexington

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02173

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/441,629

; FILING DATE: 15-MAY-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/279,217

; FILING DATE: 22-JUL-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Granahan, Patricia

; REGISTRATION NUMBER: 32,227

; REFERENCE/DOCKET NUMBER: HU95-01A

; TELEPHONE: (617) 861-6240

; TELEFAX: (617) 861-9540

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 190 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-441-629-4

Query Match

Best Local Similarity 17.2%; Score 5; DB 1; Length 190;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 INQLL 8

Db 45 INQLL 49

RESULT 35

US-08-776-207-4

; Sequence 4, Application US/08776207A

; Patent No. 6080718

; GENERAL INFORMATION:

; APPLICANT: Kirschner, Marc W.

; APPLICANT: Kinoshita, No. 6080718iyuki

; TITLE OF INVENTION: Receptor-Ligand Assay

; FILE REFERENCE: HU95-01A2

; CURRENT APPLICATION NUMBER: US/08/776,207A

;; CURRENT FILING DATE: 1997-06-23
;; EARLIER APPLICATION NUMBER: PCT/US95/09172
;; EARLIER FILING DATE: 1995-07-19
;; EARLIER APPLICATION NUMBER: 08/441,629
;; EARLIER FILING DATE: 1995-05-15
;; EARLIER APPLICATION NUMBER: 08/279,217
;; EARLIER FILING DATE: 1994-07-22
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 4
;; LENGTH: 190
;; TYPE: PRT
;; ORGANISM: Xenopus laevis
US-08-776-207-4

Query Match 17.2%; Score 5; DB 3; Length 190;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 INOLL 8
Db 45 INOLL 49

RESULT 36
PCT-US95-09172-4
;; Sequence 4, Application PC/TUS9509172
;; GENERAL INFORMATION:
;; APPLICANT: Kirschner, Marc W.
;; APPLICANT: Kinoshita, Noriyuki
;; TITLE OF INVENTION: RECEPTOR-LIGAND ASSAY
;; NUMBER OF SEQUENCES: 17
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
;; STREET: Two Militia Drive
;; CITY: Lexington
;; STATE: Massachusetts
;; COUNTRY: USA
;; ZIP: 02173
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US95/09172
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/279,217
;; FILING DATE: 22-JUL-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/441,629
;; FILING DATE: 15-MAY-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Granahan, Patricia
;; REGISTRATION NUMBER: 32,227
;; REFERENCE/DOCKET NUMBER: HU95-01A PCT
;; TELEPHONE: (617) 861-6240
;; TELEFAX: (617) 861-9540
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 190 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
PCT-US95-09172-4

Query Match 17.2%; Score 5; DB 5; Length 190;

Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 INOLL 8
Db 45 INOLL 49
RESULT 37
US-09-056-556-187
;; Sequence 187, Application US/09056556
;; Patent No. 6350456
;; GENERAL INFORMATION:
;; APPLICANT: Reed, Steven G.
;; APPLICANT: Skeiky, Yasir A.W.
;; APPLICANT: Dillon, Davin C.
;; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
;; NUMBER OF SEQUENCES: 241
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: SEED and BERRY LLP
;; STREET: 6300 Columbia Center, 701 Fifth Avenue
;; CITY: Seattle
;; STATE: Washington
;; COUNTRY: USA
;; ZIP: 98104-7092
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/056,556
;; FILING DATE: 07-APR-1998
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Makl, David J.
;; REGISTRATION NUMBER: 31,392
;; REFERENCE/DOCKET NUMBER: 210121.457
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (206) 622-4900
;; TELEFAX: (206) 682-6031
;; INFORMATION FOR SEQ ID NO: 187:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 196 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
US-09-056-556-187

Query Match 17.2%; Score 5; DB 4; Length 196;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELAPR 20
Db 45 ELAPR 49

RESULT 38
US-08-738-462-2
;; Sequence 2, Application US/08738462
;; Patent No. 5965401
;; GENERAL INFORMATION:
;; APPLICANT: Chang, Chlwen L.
;; APPLICANT: Lanier, Lewis L.
;; APPLICANT: Phillips Jr., Joseph H.
;; TITLE OF INVENTION: Purified Mammalian NK Antigens and
;; TITLE OF INVENTION: Related Reagents
;; NUMBER OF SEQUENCES: 2
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: DNAX Research Institute
;; STREET: 901 California Avenue

```
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/738,462
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/093,435
; FILING DATE: 16-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0397
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-852-9196
; TELEFAX: 415-496-1200
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 225 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-738-462-2

Query Match 17.2%; Score 5; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 LPTDS 15
Db 12 LPTDS 16

RESULT 39
PCT-US94-07587-2
; SEQUENCE 2, Application PC/TUS9407587
; GENERAL INFORMATION:
; APPLICANT: Schering Corp.
; TITLE OF INVENTION: PURIFIED MAMMALIAN NK ANTIGENS AND
; TITLE OF INVENTION: RELATED REAGENTS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering Corp.
; STREET: One Giralda Farms
; CITY: Madison
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07940
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.1
; SOFTWARE: Microsoft Word 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07587
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lunn, Paul G.
; REGISTRATION NUMBER: 32,743
; REFERENCE/DOCKET NUMBER: DX0397K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-822-7255
; TELEFAX: 201-822-7039
; INFORMATION FOR SEQ ID NO: 2:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 225 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US94-07587-2

Query Match 17.2%; Score 5; DB 5; Length 225;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 LPTDS 15
Db 12 LPTDS 16

RESULT 40
US-09-655-270A-15
; SEQUENCE 15, Application US/09655270A
; Patent No. 6329151
; GENERAL INFORMATION:
; APPLICANT: Rouviere, Pierre E.
; TITLE OF INVENTION: High Density Sampling of Differentially Expressed Prokaryotic
; FILE REFERENCE: BC1011 US NA
; CURRENT APPLICATION NUMBER: US/09/655,270A
; CURRENT FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/120,702
; PRIOR FILING DATE: 1999-February-19
; PRIOR APPLICATION NUMBER: 60/152,542
; PRIOR FILING DATE: 1999-September-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Rhodococcus erythropolis HL PM-1
; US-09-655-270A-15

Query Match 17.2%; Score 5; DB 4; Length 227;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SELAP 19
Db 86 SELAP 90

RESULT 41
US-09-651-941-17
; SEQUENCE 17, Application US/09651941
; Patent No. 6355470
; GENERAL INFORMATION:
; APPLICANT: ROUVIER, PIERRE E
; APPLICANT: WALTERS, DANA M
; APPLICANT: RAINER, RUSS
; TITLE OF INVENTION: Genes Encoding Picric Acid Degradation
; FILE REFERENCE: BC1022 US NA
; CURRENT APPLICATION NUMBER: US/09/651,941
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,545
; PRIOR FILING DATE: 1999-10-03
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Rhodococcus erythropolis HL PM-1
; US-09-651-941-17

Query Match 17.2%; Score 5; DB 4; Length 227;
```

```
;
; REFERENCE/DOCKET NUMBER: 5100-0003.00
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 269 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-447-554-5

Query Match 17.2% Score 5; DB 1; Length 269;
Best Local Similarity 100.0%; Pred. No. 2.7e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

Qy 17 LAPRS 21
Db 231 LAPRS 235

RESULT 44
US-08-468-763-17
; Sequence 17, Application US/08468763
; Patent No. 5741671
; GENERAL INFORMATION:
; APPLICANT: Agre, Peter C.
; TITLE OF INVENTION: Isolation, Cloning and Expression of
; TITLE OF INVENTION: Transmembrane Water Channel Proteins
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti
; STREET: 1001 G Street, N.W.
; CITY: Washington, D.C.
; STATE: D.C.
; COUNTRY: US
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,763
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/393,996
; FILING DATE: 24-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Posorske, Laurence H.
; REGISTRATION NUMBER: 34,698
; REFERENCE/DOCKET NUMBER: 1107.48633
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 508-9100
; TELEFAX: 202 508-9299
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 269 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-763-17

Query Match 17.2% Score 5; DB 1; Length 269;
Best Local Similarity 100.0%; Pred. No. 2.7e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

Qy 17 LAPRS 21
Db 231 LAPRS 235

US-09-364-230-34
; Sequence 34, Application US/09364230
; Patent No. 6348339
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Hitz, William D.
; APPLICANT: Kiney, Anthony J.
; APPLICANT: Rafalski, J. Antoni
; TITLE OF INVENTION: Enzymes Involved in Degradation of Branched-Chain Amino Acids
; FILE REFERENCE: BB-1178
; CURRENT FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,990
; EARLIER FILING DATE: July 31, 1998
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 34
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Triticum aestivum
; US-09-364-230-34

Query Match 17.2% Score 5; DB 4; Length 255;
Best Local Similarity 100.0%; Pred. No. 2.6e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

Qy 3 CINOL 7
Db 139 CINOL 143

RESULT 43
US-08-447-554-5
; Sequence 5, Application US/08447554
; Patent No. 5661003
; GENERAL INFORMATION:
; APPLICANT: FUSHIMI, KIYOHIKE
; APPLICANT: UCHIDA, SHINICHI
; APPLICANT: SASAKI, SEI
; APPLICANT: MARUMO, FUMIAKI
; TITLE OF INVENTION: WATER CHANNEL
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mortlison & Foerster
; STREET: 2000 Pennsylvania Ave. NW, Ste. 5500
; CITY: Washington, DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,554
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,365
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
```

```
;
; REFERENCE/DOCKET NUMBER: 5100-0003.00
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 269 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-447-554-5

Query Match 17.2% Score 5; DB 1; Length 269;
Best Local Similarity 100.0%; Pred. No. 2.7e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

Qy 17 LAPRS 21
Db 231 LAPRS 235

RESULT 44
US-08-468-763-17
; Sequence 17, Application US/08468763
; Patent No. 5741671
; GENERAL INFORMATION:
; APPLICANT: Agre, Peter C.
; TITLE OF INVENTION: Isolation, Cloning and Expression of
; TITLE OF INVENTION: Transmembrane Water Channel Proteins
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti
; STREET: 1001 G Street, N.W.
; CITY: Washington, D.C.
; STATE: D.C.
; COUNTRY: US
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,763
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/393,996
; FILING DATE: 24-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Posorske, Laurence H.
; REGISTRATION NUMBER: 34,698
; REFERENCE/DOCKET NUMBER: 1107.48633
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 508-9100
; TELEFAX: 202 508-9299
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 269 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-763-17

Query Match 17.2% Score 5; DB 1; Length 269;
Best Local Similarity 100.0%; Pred. No. 2.7e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

Qy 17 LAPRS 21
Db 231 LAPRS 235
```

Db 231 LAPRS 235

RESULT 45

US-08-448-160-5
; Sequence 5, Application US/08448160
; Patent No. 5785986
; GENERAL INFORMATION:
; APPLICANT: FUSHIMI, KIYOHIDE
; APPLICANT: UCHIDA, SHINICHI
; APPLICANT: SASAKI, SEI
; APPLICANT: MARUMO, FUMIAKI
; TITLE OF INVENTION: WATER CHANNEL
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Ave. NW, Ste. 5500
; CITY: Washington, DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,160
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,554
; FILING DATE:
; APPLICATION NUMBER: US 08/126,365
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 5100-0003.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 269 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-448-160-5

Query Match 17.2%; Score 5; DB 1; Length 269;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LAPRS 21

Db 231 LAPRS 235

RESULT 46

US-08-393-996A-17
; Sequence 17, Application US/08393996A
; Patent No. 5858702
; GENERAL INFORMATION:
; APPLICANT: Agre, Peter C.
; TITLE OF INVENTION: Isolation, Cloning and Expression of
; TITLE OF INVENTION: Transmembrane Water Channel Proteins
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti
; STREET: 1001 G Street, N.W.
; CITY: Washington, D.C.

STATE: D.C.
COUNTRY: US
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/393,996A
FILING DATE: 24-FEB-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Posorske, Laurence H.
REGISTRATION NUMBER: 34,698
REFERENCE/DOCKET NUMBER: 1107.48633
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 508-9100
TELEFAX: 202 508-9299
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 269 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-393-996A-17

Query Match 17.2%; Score 5; DB 2; Length 269;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LAPRS 21

Db 231 LAPRS 235

RESULT 47

US-09-077-675A-12
; Sequence 12, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-077-675A-12

Query Match 17.2%; Score 5; DB 4; Length 271;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LLCKL 11
 |||||
Db 19 LLCKL 23

RESULT 48
US-09-077-675A-5
; Sequence 5, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-077-675A-5

Query Match 17.2%; Score 5; DB 4; Length 289;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 7 LLCKL 11
 |||||
Db 114 LLCKL 118
RESULT 49
US-09-077-675A-10
; Sequence 10, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-077-675A-10

Query Match 17.2%; Score 5; DB 4; Length 289;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LLCKL 11
 |||||
Db 114 LLCKL 118

RESULT 50
US-08-576-626A-34
; Sequence 34, Application US/08576626A
; Patent No. 5998194
; GENERAL INFORMATION:
; APPLICANT: Summers, R.G.
; APPLICANT: Katz, L.
; APPLICANT: Donadio, S.

APPLICANT: Staver, M.J.
TITLE OF INVENTION: POLYKETIDE-ASSOCIATED SUGAR
TITLE OF INVENTION: BIOSYNTHESIS GENES
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: Illinois
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/576,626A
FILING DATE: 21-DEC-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dianne Casuto
REGISTRATION NUMBER: P-40,943
REFERENCE/DOCKET NUMBER: 5857.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (847) 938-3137
TELEFAX: (847) 938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 290 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5998194e
US-08-576-626A-34

Query Match 17.2% Score 5; DB 2; Length 290;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 INQLL 8
|
|
|
|
Db 41 INQLL 45

Search completed: August 22, 2002, 11:17:39
Job time: 46 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 11:16:53 ; Search time 52.19 Seconds
(without alignments)
61.719 Million cell updates/sec

Title: US-09-856-199-1

Perfect score: 29

Sequence: 1 KRCINOLLCKLPTDSELAPRSXCCHCRHL 29

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1200 summaries

Database : A_Geneseq_032802.*

```

1: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	72.4	21	21	Human selenoprotein
2	21	72.4	29	21	Human selenoprotein
3	21	72.4	103	21	Human selenoprotein
4	21	72.4	110	21	Human secreted protein
5	21	72.4	240	21	Human prostate cancer
6	21	72.4	381	21	Human selenoprotein
7	20	69.0	20	21	Human selenoprotein
8	6	20.7	51	22	Novel human respiratory
9	6	20.7	106	21	Arabidopsis thaliana
10	6	20.7	112	21	Arabidopsis thaliana
11	6	20.7	116	21	Arabidopsis thaliana

12	6	20.7	120	22	ABG27736	Novel human diaphanous
13	6	20.7	134	22	ABG27735	Novel human diaphanous
14	6	20.7	170	21	ABG29772	Arabidopsis thaliana
15	6	20.7	176	21	ABG29771	Arabidopsis thaliana
16	6	20.7	180	21	AAU29770	Arabidopsis thaliana
17	6	20.7	276	22	AAU02498	Murine trinucleotide repeat
18	6	20.7	311	21	ABG5022	Arabidopsis thaliana
19	6	20.7	334	21	AA19382	Amino acid sequence
20	6	20.7	359	20	AAU06628	C. elegans tyrosyl transferase
21	6	20.7	359	21	AAU84309	A. thaliana tyrosyl transferase
22	6	20.7	404	21	AA132501	S. typhimurium Mif
23	6	20.7	423	21	AAU16444	Arabidopsis thaliana
24	6	20.7	427	22	AAU19832	Human novel extracellular matrix protein
25	6	20.7	460	21	AA19384	Amino acid sequence
26	6	20.7	478	15	AA146065	Mutant alpha-amylase
27	6	20.7	478	16	AAU79025	Mature taka-amyloid
28	6	20.7	478	16	AAU72450	Aspergillus oryzae
29	6	20.7	478	16	AAU78270	Aspergillus oryzae
30	6	20.7	478	17	AAU14500	Aspergillus oryzae
31	6	20.7	489	21	AA19380	Amino acid sequence
32	6	20.7	498	22	AA184206	Amino acid sequence
33	6	20.7	503	22	AB167817	Drosophila melanogaster
34	6	20.7	556	21	AA19381	Amino acid sequence
35	6	20.7	602	22	AB171195	Drosophila melanogaster
36	6	20.7	615	21	AA19383	Amino acid sequence
37	6	20.7	616	17	AAU86867	Rat protocadherin
38	6	20.7	645	21	AA19391	Amino acid sequence
39	6	20.7	704	22	AB164989	Drosophila melanogaster
40	6	20.7	704	22	AB166726	Drosophila melanogaster
41	6	20.7	711	21	AA19379	Amino acid sequence
42	6	20.7	734	22	AB169061	Drosophila melanogaster
43	6	20.7	771	21	AA19392	Amino acid sequence
44	6	20.7	791	22	AAU19687	Human novel extracellular matrix protein
45	6	20.7	816	22	AB168766	Drosophila melanogaster
46	6	20.7	830	22	AB162835	Drosophila melanogaster
47	6	20.7	832	22	ABG23820	Novel human diaphanous
48	6	20.7	858	22	AAU79446	Human protein sequence
49	6	20.7	867	21	AA19390	Amino acid sequence
50	6	20.7	916	20	AAU14743	Human PRO707 protease
51	6	20.7	916	21	AA144299	Human PRO707 (UNC)
52	6	20.7	916	22	AAU29062	Human PRO polypeptide
53	6	20.7	931	22	AAU78649	Human protein sequence
54	6	20.7	932	22	ABG17156	Novel human diaphanous
55	6	20.7	949	22	AB12315	Human protocadherin
56	6	20.7	949	22	AAU79633	Human protein sequence
57	6	20.7	1042	22	AAU15096	Protein encoded by Drosophila melanogaster
58	6	20.7	1333	22	AB171827	Drosophila melanogaster
59	6	20.7	1779	22	AB171649	Drosophila melanogaster
60	5	17.2	14	22	AAU69182	Human Acetylcholinesterase
61	5	17.2	15	18	AAU38998	Peptide resembling human secreted alpha
62	5	17.2	15	22	AA160189	Human secreted alpha
63	5	17.2	15	22	AA160193	Murine secreted alpha
64	5	17.2	19	22	AA185579	Human PCPLD cDNA
65	5	17.2	21	22	AB141543	Peptide #9049 enco
66	5	17.2	21	22	AAU62414	Human brain expressed protein
67	5	17.2	21	22	AAU75222	Human bone marrow
68	5	17.2	21	22	AAU35334	Peptide #9371 enco
69	5	17.2	23	20	AAU78250	Fragment of human
70	5	17.2	29	22	AA186267	Human immune/haema
71	5	17.2	30	22	AB141859	Peptide #9365 enco
72	5	17.2	30	22	AB125560	Protein #7559 enco
73	5	17.2	30	22	AAU62732	Human brain expressed
74	5	17.2	30	22	AAU75551	Human bone marrow
75	5	17.2	30	22	AAU20633	Peptide #7067 enco
76	5	17.2	30	22	AAU35658	Peptide #9695 enco
77	5	17.2	34	21	AAU76244	Fragment of human
78	5	17.2	35	21	AAU89425	Core polypeptide f
79	5	17.2	35	22	AB100833	Viral Dp178/107-11
80	5	17.2	35	22	AB123038	Viral core polypep
81	5	17.2	35	22	AAU13379	Dp178-like Dp107-1
82	5	17.2	35	22	AAU77826	Core polypeptide T
83	5	17.2	37	22	AAU64874	Human 5' EST relat
84	5	17.2	39	22	AB143655	Peptide #11161 enc

85	5	17.2	39	22	ABB26606	Protein #8605 enco	158	5	17.2	66	22	ABBI8941	Protein #940 enco
86	5	17.2	39	22	AAM64603	Human brain expres	159	5	17.2	66	22	AAM96385	Human reproductiv
87	5	17.2	39	22	AAM77394	Human bone marrow	160	5	17.2	66	22	AAM54262	Human brain expres
88	5	17.2	39	22	AAM21322	Peptide #7756 enco	161	5	17.2	66	22	AAM66656	Human bone marrow
89	5	17.2	39	22	AAM37549	Peptide #11586 enc	162	5	17.2	66	22	AAM14524	Peptide #958 enco
90	5	17.2	40	21	AAB38468	Fragment of human	163	5	17.2	66	22	AAM26941	Peptide #978 enco
91	5	17.2	40	21	AAB13302	Caenorhabditis ele	164	5	17.2	66	22	AAM02250	Peptide #932 enco
92	5	17.2	41	22	AAM22896	Human prostate can	165	5	17.2	68	22	AAM46787	Propionibacterium
93	5	17.2	43	22	ABB30125	Peptide #2776 enco	166	5	17.2	68	22	AAM61902	Propionibacterium
94	5	17.2	43	22	ABB35297	Peptide #2803 enco	167	5	17.2	68	22	AAM81411	Human haematologic
95	5	17.2	43	22	ABB20739	Protein #2738 enco	168	5	17.2	69	22	ABB69294	Drosophila melanog
96	5	17.2	43	22	AAM56126	Human brain expres	169	5	17.2	69	22	AAG74320	Human colon cancer
97	5	17.2	43	22	AAM68498	Human bone marrow	170	5	17.2	70	22	ABG14714	Novel human diagno
98	5	17.2	43	22	AAM16304	Peptide #2738 enco	171	5	17.2	71	22	AAG46697	Propionibacterium
99	5	17.2	43	22	AAM28797	Peptide #2834 enco	172	5	17.2	72	22	ABG25517	Novel human diagno
100	5	17.2	43	22	AAM04040	Peptide #2722 enco	173	5	17.2	72	22	AAG62693	Human signal recog
101	5	17.2	44	22	ABB43929	Peptide #11435 enc	174	5	17.2	73	22	AAM49937	Propionibacterium
102	5	17.2	44	22	ABB26844	Protein #8843 enco	175	5	17.2	74	22	AAM85074	Human immune/haema
103	5	17.2	44	22	AAM64919	Human brain expres	176	5	17.2	75	22	AAG75382	Human colon cancer
104	5	17.2	44	22	AAM77655	Human bone marrow	177	5	17.2	76	22	AAM42947	Propionibacterium
105	5	17.2	44	22	AAM21576	Peptide #8010 enco	178	5	17.2	77	22	AAM81901	Human haematologic
106	5	17.2	44	22	AAM37855	Peptide #11892 enc	179	5	17.2	79	22	AAU57570	Propionibacterium
107	5	17.2	49	22	AAM40097	Human polypeptide	180	5	17.2	80	22	AAU56773	Propionibacterium
108	5	17.2	50	22	AAM85733	Human immune/haema	181	5	17.2	80	22	ABG24340	Novel human diagno
109	5	17.2	51	20	AAU13100	Human secreted pro	182	5	17.2	81	22	ABB30684	Peptide #3335 enco
110	5	17.2	51	22	AAU66626	Propionibacterium	183	5	17.2	81	22	ABB35857	Peptide #3363 enco
111	5	17.2	52	21	AAB16817	Bacteriophage Dp-1	184	5	17.2	81	22	ABB21269	Protein #3268 enco
112	5	17.2	52	22	ABB39087	Peptide #6593 enco	185	5	17.2	81	22	AAM56663	Human brain expres
113	5	17.2	52	22	ABB24003	Protein #6002 enco	186	5	17.2	81	22	AAM69039	Human bone marrow
114	5	17.2	52	22	AAM59739	Human brain expres	187	5	17.2	81	22	AAM16875	Peptide #3309 enco
115	5	17.2	52	22	AAM72321	Human bone marrow	188	5	17.2	81	22	AAM29359	Peptide #3396 enco
116	5	17.2	52	22	AAM23580	Peptide #6617 enco	189	5	17.2	81	22	AAU04577	Peptide #3259 enco
117	5	17.2	54	22	AAU61270	Propionibacterium	190	5	17.2	82	22	AAU11994	Human polypeptide
118	5	17.2	54	22	AAM65645	Human brain expres	191	5	17.2	83	21	AAG21733	Arabidopsis thalia
119	5	17.2	54	22	AAM22093	Peptide #8527 enco	192	5	17.2	83	21	AAG23682	Arabidopsis thalia
120	5	17.2	56	22	ABB28752	Peptide #1403 enco	193	5	17.2	83	21	AAG25178	Arabidopsis thalia
121	5	17.2	56	22	ABB29025	Peptide #1676 enco	194	5	17.2	83	21	AAG25183	Arabidopsis thalia
122	5	17.2	56	22	ABB33939	Peptide #1445 enco	195	5	17.2	83	21	AAG26691	Arabidopsis thalia
123	5	17.2	56	22	ABB34189	Peptide #1695 enco	196	5	17.2	83	21	AAG47677	Arabidopsis thalia
124	5	17.2	56	22	ABB19375	Protein #1374 enco	197	5	17.2	83	22	ABG04574	Novel human diagno
125	5	17.2	56	22	AAM54702	Human brain expres	198	5	17.2	83	22	ABG13645	Novel human diagno
126	5	17.2	56	22	AAM67102	Human bone marrow	199	5	17.2	83	22	ABG13649	Novel human diagno
127	5	17.2	56	22	AAM14963	Peptide #1397 enco	200	5	17.2	84	21	AAG28671	Arabidopsis thalia
128	5	17.2	56	22	AAM15197	Peptide #1631 enco	201	5	17.2	84	22	ABB60706	Drosophila melanog
129	5	17.2	56	22	AAM27397	Peptide #1434 enco	202	5	17.2	84	22	ABBI7314	Human nervous syst
130	5	17.2	56	22	AAM202689	Peptide #1371 enco	203	5	17.2	84	22	AAM83423	Human immune/haema
131	5	17.2	56	22	AAM02937	Peptide #1619 enco	204	5	17.2	84	22	AAU17033	Human novel secret
132	5	17.2	57	22	AAB60199	Human secreted alp	205	5	17.2	86	22	AAU74009	Human bone marrow
133	5	17.2	57	22	AAB62742	Murine secreted al	206	5	17.2	87	20	AAU01444	Secreted protein e
134	5	17.2	58	21	AAB13275	Caenorhabditis ele	207	5	17.2	89	21	AAG54666	Arabidopsis thalia
135	5	17.2	59	20	AAU19615	SEQ ID NO 333 from	208	5	17.2	89	22	AAU58860	Propionibacterium
136	5	17.2	59	22	AAM53459	Propionibacterium	209	5	17.2	90	21	AAG58765	Arabidopsis thalia
137	5	17.2	61	19	AAU75196	Human secreted pro	210	5	17.2	90	22	AAU47327	Propionibacterium
138	5	17.2	61	21	AAG28672	Arabidopsis thalia	211	5	17.2	90	22	AAU13205	Human polypeptide
139	5	17.2	62	22	ABB44343	Peptide #11849 enc	212	5	17.2	91	22	AAU61334	Propionibacterium
140	5	17.2	62	22	ABB27202	Protein #9201 enco	213	5	17.2	92	22	AAU59833	Propionibacterium
141	5	17.2	62	22	AAM65393	Human brain expres	214	5	17.2	94	20	AAU34822	Chlamydia pneumoni
142	5	17.2	62	22	AAM78084	Human bone marrow	215	5	17.2	95	22	AAU45566	Propionibacterium
143	5	17.2	62	22	AAM21967	Peptide #8401 enco	216	5	17.2	95	22	ABG14223	Novel human diagno
144	5	17.2	62	22	AAM38307	Peptide #12344 enc	217	5	17.2	95	22	ABB20508	Protein #2507 enco
145	5	17.2	63	21	AAG12614	Zea mays protein f	218	5	17.2	95	22	AAM55911	Human brain expres
146	5	17.2	63	21	AAG03312	Human secreted pro	219	5	17.2	95	22	AAM68281	Human bone marrow
147	5	17.2	64	22	ABG24994	Novel human diagno	220	5	17.2	95	22	AAM16104	Peptide #2538 enco
148	5	17.2	65	22	AAU44061	Propionibacterium	221	5	17.2	95	22	AAM28595	Peptide #2632 enco
149	5	17.2	65	22	ABU42035	Peptide #9541 enco	222	5	17.2	95	22	AAM03831	Peptide #2513 enco
150	5	17.2	65	22	ABB25649	Protein #7648 enco	223	5	17.2	96	21	AAU41677	Human ORFX ORF1441
151	5	17.2	65	22	AAM62916	Human brain expres	224	5	17.2	96	22	AAU57248	Propionibacterium
152	5	17.2	65	22	AAM75728	Human bone marrow	225	5	17.2	97	21	AAG27369	Arabidopsis thalia
153	5	17.2	65	22	AAM35837	Peptide #9874 enco	226	5	17.2	97	22	AAU08492	Mouse VMGLOM short
154	5	17.2	66	20	AAU48333	Human prostate can	227	5	17.2	97	22	AAG76566	Human colon cancer
155	5	17.2	66	22	AAU59642	Propionibacterium	228	5	17.2	98	21	AAB41005	Human ORFX ORF769
156	5	17.2	66	22	ABB28305	Human peptide #956	229	5	17.2	99	21	AAG19943	Arabidopsis thalia
157	5	17.2	66	22	ABB33481	Peptide #987 enco	230	5	17.2	100	22	AAU65486	Propionibacterium

231	5	17.2	101	20	AA929071	T. gondii immunoge	304	5	17.2	152	21	AA913984	Arabidopsis thalia
232	5	17.2	101	20	AA929055	T. gondii immunoge	305	5	17.2	152	21	AA976009	Murine TGF-alpha h
233	5	17.2	101	20	AA950222	Human clone vb14_1	306	5	17.2	152	22	AB669865	Drosophila melanog
234	5	17.2	101	22	AA064310	Propionibacterium	307	5	17.2	152	22	AA655948	Skin cell protein,
235	5	17.2	101	22	AAU25526	T. gondii immunoge	308	5	17.2	153	22	AAU65026	Propionibacterium
236	5	17.2	101	22	AAU25542	T. gondii immunoge	309	5	17.2	154	21	AA929507	Arabidopsis thalia
237	5	17.2	101	22	AA94985	Human protein sequ	310	5	17.2	154	21	AA949549	Arabidopsis thalia
238	5	17.2	102	22	AA973809	Human colon cancer	311	5	17.2	154	22	AAU05052	Propionibacterium
239	5	17.2	103	21	AA922463	Arabidopsis thalia	312	5	17.2	155	21	AA933217	Pinus radiata tran
240	5	17.2	103	21	AA922481	Arabidopsis thalia	313	5	17.2	155	21	AA924395	Arabidopsis thalia
241	5	17.2	104	22	AB603599	Human musculoskele	314	5	17.2	155	22	AAO00013	Human polypeptide
242	5	17.2	104	22	AAU23296	Novel human enzyme	315	5	17.2	156	21	AA912292	Zea mays protein f
243	5	17.2	106	21	AA936418	Arabidopsis thalia	316	5	17.2	157	21	AA918821	Zea mays protein f
244	5	17.2	107	22	AB615760	Human nervous syst	317	5	17.2	157	22	ABG12494	Novel human diagno
245	5	17.2	108	21	AA955584	Arabidopsis thalia	318	5	17.2	158	21	AA908047	Arabidopsis thalia
246	5	17.2	108	22	AAU16435	Human novel secret	319	5	17.2	158	21	AA954562	Zea mays protein f
247	5	17.2	109	21	AA943177	Human ORF2941	320	5	17.2	158	21	AA976112	Rat integrin homol
248	5	17.2	109	22	AA013249	Human polypeptide	321	5	17.2	158	22	AAU30609	Novel human secret
249	5	17.2	110	21	AA919826	Arabidopsis thalia	322	5	17.2	158	22	AA656051	Skin cell protein,
250	5	17.2	110	21	AA934827	Arabidopsis thalia	323	5	17.2	159	21	AA911630	Arabidopsis thalia
251	5	17.2	112	21	AA908935	Arabidopsis thalia	324	5	17.2	159	21	AA951644	Arabidopsis thalia
252	5	17.2	112	21	AA927368	Arabidopsis thalia	325	5	17.2	159	22	AAU30631	Novel human secret
253	5	17.2	112	21	AA937290	Arabidopsis thalia	326	5	17.2	160	20	AA935584	Chlamydia pneumoni
254	5	17.2	112	22	ABG14578	Novel human diagno	327	5	17.2	160	22	AB658297	Drosophila melanog
255	5	17.2	113	21	AA915828	Arabidopsis thalia	328	5	17.2	162	22	AAU30297	Novel human secret
256	5	17.2	113	21	AA947805	Arabidopsis thalia	329	5	17.2	162	22	AAU30297	Novel human secret
257	5	17.2	113	21	AA947808	Arabidopsis thalia	330	5	17.2	163	22	AAU55844	Propionibacterium
258	5	17.2	113	22	AA013781	Human polypeptide	331	5	17.2	164	21	AA914907	Arabidopsis thalia
259	5	17.2	113	22	AA974439	Human colon cancer	332	5	17.2	164	21	AA922809	Arabidopsis thalia
260	5	17.2	113	22	AA939393	Human protein sequ	333	5	17.2	164	21	AA951649	Arabidopsis thalia
261	5	17.2	114	22	ABG12492	Novel human diagno	334	5	17.2	165	22	AA987469	Human immune/haema
262	5	17.2	116	22	ABG28581	Novel human diagno	335	5	17.2	166	22	AA975568	Human secreted pro
263	5	17.2	117	17	AA961135	Full length Defens	336	5	17.2	168	22	AAU30665	Novel human secret
264	5	17.2	117	22	AA90505	C glutamicum prote	337	5	17.2	168	22	AA98979	E. coli growth and
265	5	17.2	119	22	ABG19689	Novel human diagno	338	5	17.2	170	22	AA980935	Human nPCR12. Ho
266	5	17.2	120	20	AA948601	Human breast tumou	339	5	17.2	173	20	AA936238	Human secreted pro
267	5	17.2	121	20	AA936417	Arabidopsis thalia	340	5	17.2	173	22	AA972620	Human angiotensin
268	5	17.2	121	21	AA955583	Arabidopsis thalia	341	5	17.2	174	22	AA660184	Human secreted alp
269	5	17.2	122	17	AA989746	Art-1 interacting	342	5	17.2	174	22	AA660186	Murine secreted alp
270	5	17.2	123	22	AA94901	Human reproductive	343	5	17.2	175	22	AAU23698	Novel human enzyme
271	5	17.2	125	21	AA915827	Arabidopsis thalia	344	5	17.2	178	21	AA943774	Arabidopsis thalia
272	5	17.2	125	21	AA947804	Arabidopsis thalia	345	5	17.2	178	22	AA988233	Human immune/haema
273	5	17.2	125	22	AAO00492	Human polypeptide	346	5	17.2	179	22	AB632312	Peptide #4963 enco
274	5	17.2	126	21	AA936416	Arabidopsis thalia	347	5	17.2	179	22	AB637574	Peptide #5080 enco
275	5	17.2	127	21	AA954563	Zea mays protein f	348	5	17.2	179	22	AA958228	Human brain expres
276	5	17.2	127	22	AA960591	Human brain expres	349	5	17.2	179	22	AA970682	Human bone marrow
277	5	17.2	127	22	AA973259	Human bone marrow	350	5	17.2	179	22	AA918537	Peptide #4971 enco
278	5	17.2	127	22	AA984160	Human immune/haema	351	5	17.2	179	22	AA930993	Peptide #5030 enco
279	5	17.2	127	22	AA933460	Peptide #7497 enco	352	5	17.2	179	22	AA941952	Human polypeptide
280	5	17.2	128	22	AAU57766	Propionibacterium	353	5	17.2	180	21	AA943773	Arabidopsis thalia
281	5	17.2	129	21	AA963227	Gene 45 human secr	354	5	17.2	182	22	ABG16552	Novel human diagno
282	5	17.2	130	21	AA932902	Arabidopsis thalia	355	5	17.2	184	21	AA925847	Arabidopsis thalia
283	5	17.2	130	22	AB651174	Human secreted pro	356	5	17.2	184	22	AA93431	Human protein sequ
284	5	17.2	130	22	AA985139	Human immune/haema	357	5	17.2	185	21	AA914906	Arabidopsis thalia
285	5	17.2	131	21	AA901409	Fragment of macaqu	358	5	17.2	188	22	AA988879	Human immune/haema
286	5	17.2	132	21	AA924397	Arabidopsis thalia	359	5	17.2	190	17	AA990768	FGF receptor ligan
287	5	17.2	133	21	AA901145	Human secreted pro	360	5	17.2	194	22	AAU62234	Propionibacterium
288	5	17.2	133	22	ABG01793	Novel human diagno	361	5	17.2	196	19	AA981735	M. tuberculosis im
289	5	17.2	134	21	AA958888	Breast and ovarian	362	5	17.2	196	19	AA964370	Mycobacterium tube
290	5	17.2	138	21	AA922810	Arabidopsis thalia	363	5	17.2	196	20	AA939167	M. tuberculosis an
291	5	17.2	138	21	AA924396	Arabidopsis thalia	364	5	17.2	196	20	AA939024	M. tuberculosis re
292	5	17.2	139	21	AA951650	Arabidopsis thalia	365	5	17.2	196	22	AAU59416	Propionibacterium
293	5	17.2	141	21	AA953681	Human colon cancer	366	5	17.2	196	22	AA973614	Human colon cancer
294	5	17.2	142	22	AB663681	Drosophila melanog	367	5	17.2	197	21	AA955601	Arabidopsis thalia
295	5	17.2	143	22	AB688845	Drosophila melanog	368	5	17.2	197	22	ABG00149	Novel human diagno
296	5	17.2	143	22	ABG13842	Novel human diagno	369	5	17.2	197	22	AAU30103	Novel human secret
297	5	17.2	144	22	AA994547	Human reproductive	370	5	17.2	199	22	AA925846	Arabidopsis thalia
298	5	17.2	146	22	AB611229	Human Leu-rich pro	371	5	17.2	199	22	AAU27704	Human full-length
299	5	17.2	147	22	AAU30905	Novel human secret	372	5	17.2	199	22	AAU30364	Novel human secret
300	5	17.2	150	21	AA901131	Human secreted pro	373	5	17.2	199	22	AA960183	Human secreted alp
301	5	17.2	150	22	AAU55896	Propionibacterium	374	5	17.2	199	22	AA960185	Murine secreted al
302	5	17.2	150	22	AA931336	Human protein sequ	375	5	17.2	200	21	AA934426	Arabidopsis thalia
303	5	17.2	151	22	AB612301	Human ketosephosph	376	5	17.2	200	22	AAU30430	Novel human secret

377	5	17.2	202	21	AAG11629	Arabidopsis thalia	450	5	17.2	246	22	ARG07129	Novel human diagno
378	5	17.2	202	21	AAG51643	Arabidopsis thalia	451	5	17.2	246	22	ARG26936	Novel human diagno
379	5	17.2	203	22	AAU20547	Human secreted pro	452	5	17.2	246	22	AM99917	Human polypeptide
380	5	17.2	205	21	AGA33772	Arabidopsis thalia	453	5	17.2	248	22	ABG24603	Novel human diagno
381	5	17.2	205	22	AAU25697	G protein-coupled	454	5	17.2	249	17	AA97263	Herpesvirus saimir
382	5	17.2	205	22	AAU25815	Human protein sequ	455	5	17.2	250	22	ABG00213	Novel human diagno
383	5	17.2	206	21	AAU09952	Arabidopsis thalia	456	5	17.2	250	22	ABG07120	Novel human diagno
384	5	17.2	206	21	AAU14905	Arabidopsis thalia	457	5	17.2	250	22	ABG50476	Listeria sp. strai
385	5	17.2	208	21	AGG30052	Arabidopsis thalia	458	5	17.2	252	21	AAG09516	Arabidopsis thalia
386	5	17.2	208	22	ABG11773	Novel human diagno	459	5	17.2	252	21	AAG14627	Arabidopsis thalia
387	5	17.2	208	22	AA94757	Human protein sequ	460	5	17.2	252	21	AG34425	Arabidopsis thalia
388	5	17.2	209	22	AB868554	Drosophila melanog	461	5	17.2	252	21	AG48074	Arabidopsis thalia
389	5	17.2	209	22	ABG04149	Novel human diagno	462	5	17.2	252	22	AA90595	Human secreted pro
390	5	17.2	211	21	AB384466	Fragment of human	463	5	17.2	254	21	AG09461	Arabidopsis thalia
391	5	17.2	211	22	ABG24589	Novel human diagno	464	5	17.2	254	21	AG43317	Arabidopsis thalia
392	5	17.2	212	21	AA660089	Arabidopsis thalia	465	5	17.2	255	22	ABG13404	Novel human diagno
393	5	17.2	212	22	ABG11748	Novel human diagno	466	5	17.2	255	22	AM79889	Human protein SEQ
394	5	17.2	212	22	AAU33289	Novel human secret	467	5	17.2	256	22	ABG20789	Novel human diagno
395	5	17.2	215	21	AA645458	Arabidopsis thalia	468	5	17.2	256	22	AA000765	Human bone marrow
396	5	17.2	216	21	AA805946	Protein deduced fr	469	5	17.2	256	22	AA88382	Human membrane or
397	5	17.2	217	20	AA929470	Human vanilloid re	470	5	17.2	257	17	AA94460	Uroporphyrinogen I
398	5	17.2	217	21	AA32520	Arabidopsis thalia	471	5	17.2	258	21	AA655600	Arabidopsis thalia
399	5	17.2	219	22	AA89780	C glutamicum prote	472	5	17.2	258	22	AG90666	C glutamicum prote
400	5	17.2	220	21	AAU4584	Arabidopsis thalia	473	5	17.2	258	22	AA70161	DNA encoding human
401	5	17.2	220	21	AA45443	Arabidopsis thalia	474	5	17.2	259	21	AA655599	Arabidopsis thalia
402	5	17.2	221	21	AA45457	Arabidopsis thalia	475	5	17.2	259	22	ABG24545	Novel human diagno
403	5	17.2	221	22	ABG24565	Novel human diagno	476	5	17.2	260	22	ABG61666	Drosophila melanog
404	5	17.2	222	21	AA09951	Arabidopsis thalia	477	5	17.2	260	22	ABG26923	Novel human diagno
405	5	17.2	223	21	AA32519	Arabidopsis thalia	478	5	17.2	261	22	ABG65619	Drosophila melanog
406	5	17.2	224	22	AB858988	Arabidopsis thalia	479	5	17.2	267	21	AA09515	Arabidopsis thalia
407	5	17.2	225	16	AA865189	Drosophila melanog	480	5	17.2	267	21	AAU4626	Arabidopsis thalia
408	5	17.2	225	21	AA87337	NK cell antigen DX	481	5	17.2	267	21	AA48073	Arabidopsis thalia
409	5	17.2	226	21	AA645483	Human signal pepti	482	5	17.2	267	22	AAU00484	Arabidopsis thalia
410	5	17.2	226	21	AA45443	Arabidopsis thalia	483	5	17.2	268	22	AAU33169	S. galliaeus aklev
411	5	17.2	226	22	AA90170	C glutamicum prote	484	5	17.2	269	19	AA55786	Novel human secret
412	5	17.2	227	21	AA802085	F420/NADPH oxidore	485	5	17.2	269	20	AA94322	Human aquaporin-1
413	5	17.2	229	20	AAU27341	Group B streptococ	486	5	17.2	269	22	ABG17234	Human human diagno
414	5	17.2	229	21	AA877721	Arabidopsis thalia	487	5	17.2	271	18	AAU19612	Human growth hormo
415	5	17.2	230	21	AA32942	Pinus radiata tran	488	5	17.2	271	18	AAU19219	Human growth hormo
416	5	17.2	230	22	ABG15066	Novel human diagno	489	5	17.2	271	22	AAU15537	Human immune syste
417	5	17.2	231	21	AA627720	Arabidopsis thalia	490	5	17.2	273	22	AA92777	Human protein sequ
418	5	17.2	231	21	AA30051	Arabidopsis thalia	491	5	17.2	276	21	AA31704	Arabidopsis thalia
419	5	17.2	232	21	AA627719	Arabidopsis thalia	492	5	17.2	277	22	ABG15834	Novel human diagno
420	5	17.2	233	22	ABG12360	Novel human diagno	493	5	17.2	278	21	AAU96729	PRO1800, a Hep27 h
421	5	17.2	234	22	AAU16961	Human novel secret	494	5	17.2	278	21	AAU68735	Short chain alcoh
422	5	17.2	235	22	AAU51230	Thale cress Ntr1 r	495	5	17.2	278	22	AA93414	Human protein sequ
423	5	17.2	236	22	AA42101	Human polypeptide	496	5	17.2	280	22	ABG21093	Novel human diagno
424	5	17.2	237	22	ABG69862	Drosophila melanog	497	5	17.2	280	22	ABG25564	Novel human diagno
425	5	17.2	238	21	AAU24023	Human PRO1410 prot	498	5	17.2	280	22	AAU19848	Mycobacterium tube
426	5	17.2	238	21	AA333455	Human PRO1410 prot	499	5	17.2	281	21	AAU91070	Streptomyces nogal
427	5	17.2	238	21	AA801408	Macaque TANGO 232	500	5	17.2	281	22	ABG59929	Drosophila melanog
428	5	17.2	238	21	AA801417	Human TANGO 232	501	5	17.2	281	22	AA65655	Novel protein kina
429	5	17.2	238	21	AA645456	Arabidopsis thalia	502	5	17.2	282	22	AAU33193	Novel human secret
430	5	17.2	238	21	AAU99415	Human PRO1410 (UNQ	503	5	17.2	283	22	AAU08658	Human NOV2 protein
431	5	17.2	238	22	AA666164	Protein of the inv	504	5	17.2	284	20	AAU40583	Partial amino acid
432	5	17.2	238	22	AA650967	Human PRO1410 prot	505	5	17.2	284	20	AAU85102	Thyroid hormone re
433	5	17.2	239	21	AAU09517	Arabidopsis thalia	506	5	17.2	285	21	AA339254	Human secreted pro
434	5	17.2	239	21	AAU14628	Arabidopsis thalia	507	5	17.2	285	22	ABG24540	Novel human diagno
435	5	17.2	239	21	AA648075	Arabidopsis thalia	508	5	17.2	285	22	AA63268	Human breast cance
436	5	17.2	239	22	ABG05020	Novel human diagno	509	5	17.2	287	15	AAU33544	Thyroid hormone re
437	5	17.2	240	21	AAU32518	Arabidopsis thalia	510	5	17.2	287	20	AAU92390	Human TR-interacti
438	5	17.2	240	22	AB36976	Peptide #4482 enco	511	5	17.2	287	21	AAU09460	Arabidopsis thalia
439	5	17.2	240	22	AAU23158	Novel human enzyme	512	5	17.2	287	21	AAU43316	Arabidopsis thalia
440	5	17.2	240	22	AAU57709	Human brain expres	513	5	17.2	287	21	AAU77933	A. thaliana enviro
441	5	17.2	240	22	AAU70118	Human bone marrow	514	5	17.2	287	22	ABG01402	Novel human diagno
442	5	17.2	240	22	AAU30457	Peptide #4494 enco	515	5	17.2	287	22	ABG02010	Novel human diagno
443	5	17.2	241	20	AAU37435	Amino acid sequenc	516	5	17.2	287	22	ABG03073	Novel human diagno
444	5	17.2	242	18	AAU10017	G-protein coupled	517	5	17.2	287	22	ABG10191	Novel human diagno
445	5	17.2	242	18	AAU20550	H. pylori cytoplas	518	5	17.2	287	22	ABG12925	Novel human diagno
446	5	17.2	243	22	ABG07320	Novel human diagno	519	5	17.2	287	22	ABG15688	Novel human diagno
447	5	17.2	244	15	AAU52829	HTP-reductase. Sa	520	5	17.2	287	22	ABG19797	Novel human diagno
448	5	17.2	244	21	AAU10110	Yeast 2,5-diamino-	521	5	17.2	287	22	ABG22800	Novel human diagno
449	5	17.2	245	22	AAU1883	Human polypeptide	522	5	17.2	287	22	ABG23933	Novel human diagno

523	5	17.2	288	21	AAB52594	Helicobacter pylori	596	5	17.2	318	22	AAU19243	Human G protein-co
524	5	17.2	288	21	AAG46835	Arabidopsis thalia	597	5	17.2	319	10	AAU19243	Polyepitide with l
525	5	17.2	288	22	ABG09552	Novel human diago	598	5	17.2	319	17	AAU19243	Pseudomonas glumae
526	5	17.2	288	22	ABG09552	Escherichia coli p	599	5	17.2	319	17	AAU19243	Pseudomonas glumae
527	5	17.2	289	18	AAU19609	Pig growth hormone	600	5	17.2	319	17	AAU19609	Mature Pseudomonas
528	5	17.2	289	18	AAU19611	Human growth hormo	601	5	17.2	319	17	AAU19611	Mature Pseudomonas
529	5	17.2	289	18	AAU19216	Swine growth hormo	602	5	17.2	319	17	AAU19216	Mature Pseudomonas
530	5	17.2	289	18	AAU19218	Human growth hormo	603	5	17.2	319	17	AAU19218	Mature Pseudomonas
531	5	17.2	290	21	AAU19218	Human ORFX ORF2004	604	5	17.2	319	17	AAU19218	Mature Pseudomonas
532	5	17.2	290	21	AAU19218	Arabidopsis thalia	605	5	17.2	319	17	AAU19218	Mature Pseudomonas
533	5	17.2	290	21	AAU19218	Arabidopsis thalia	606	5	17.2	319	17	AAU19218	Mature Pseudomonas
534	5	17.2	290	21	AAU19218	Arabidopsis thalia	607	5	17.2	319	17	AAU19218	Mature Pseudomonas
535	5	17.2	291	21	AAU19218	Arabidopsis thalia	608	5	17.2	319	17	AAU19218	Mature Pseudomonas
536	5	17.2	291	22	ABG04645	Novel human diago	609	5	17.2	319	22	ABG04645	Mature Pseudomonas
537	5	17.2	292	22	ABG15091	Novel human diago	610	5	17.2	320	15	AAU19218	C glutamic prote
538	5	17.2	293	13	AAU19218	Envelope proteins	611	5	17.2	320	15	AAU19218	Pseudomonas deri
539	5	17.2	293	13	AAU19218	Human CAAX process	612	5	17.2	320	17	AAU19218	Pseudomonas cepaci
540	5	17.2	294	21	AAU19218	Arabidopsis thalia	613	5	17.2	320	17	AAU19218	Pseudomonas cepaci
541	5	17.2	294	21	AAU19218	Arabidopsis thalia	614	5	17.2	320	17	AAU19218	Pseudomonas cepaci
542	5	17.2	294	21	AAU19218	Arabidopsis thalia	615	5	17.2	320	21	AAU19218	Arabidopsis thalia
543	5	17.2	294	21	AAU19218	Arabidopsis thalia	616	5	17.2	322	22	AAU19218	Arabidopsis thalia
544	5	17.2	294	21	AAU19218	Arabidopsis thalia	617	5	17.2	322	22	AAU19218	S. epidermidis ope
545	5	17.2	296	21	AAU19218	Arabidopsis thalia	618	5	17.2	322	22	AAU19218	S. epidermidis ope
546	5	17.2	298	22	AAU19218	Propionibacterium	619	5	17.2	323	22	AAU19218	Novel human diago
547	5	17.2	298	22	AAU19218	Human NOV1 protein	620	5	17.2	324	21	AAU19218	Arabidopsis thalia
548	5	17.2	298	22	AAU19218	Human NOV3 protein	621	5	17.2	324	22	AAU19218	Novel human diago
549	5	17.2	299	21	AAU19218	Arabidopsis thalia	622	5	17.2	324	22	AAU19218	Human olfactory re
550	5	17.2	299	21	AAU19218	Arabidopsis thalia	623	5	17.2	325	14	AAU19218	ced-4 gene II prod
551	5	17.2	300	22	ABG20320	Novel human diago	624	5	17.2	327	17	AAU19218	Hepatitis E virus
552	5	17.2	300	22	ABG20320	Novel human diago	625	5	17.2	327	18	AAU19218	Hepatitis E virus
553	5	17.2	300	22	ABG22452	Novel human diago	626	5	17.2	327	20	AAU19218	Fragment of human
554	5	17.2	301	22	AAU19218	Klebsiella pneumonia	627	5	17.2	328	17	AAU19218	HCW Toledo strain
555	5	17.2	301	22	AAU19218	Human polypeptide	628	5	17.2	329	20	AAU19218	Human farnesylated
556	5	17.2	302	14	AAU19218	LACI. Rattus ratt	629	5	17.2	329	20	AAU19218	Mouse CAAX process
557	5	17.2	302	17	AAU19218	Lipoprotein-associ	630	5	17.2	329	20	AAU19218	Human RCE1 (hRCE1)
558	5	17.2	302	22	ABG24511	Novel human diago	631	5	17.2	329	21	AAU19218	Human RCE1 (farnes)
559	5	17.2	302	22	ABG24511	Putative P. abyss	632	5	17.2	329	21	AAU19218	Mouse RCE1 (farnes)
560	5	17.2	303	11	AAU19218	Intracellular alka	633	5	17.2	329	22	AAU19218	Novel human diago
561	5	17.2	304	21	AAU19218	Zea mays protein f	634	5	17.2	329	22	AAU19218	Human ras converti
562	5	17.2	307	22	ABG09476	Novel human diago	635	5	17.2	330	21	AAU19218	Arabidopsis thalia
563	5	17.2	309	21	AAU19218	Arabidopsis thalia	636	5	17.2	331	22	AAU19218	Novel human diago
564	5	17.2	309	21	AAU19218	Arabidopsis thalia	637	5	17.2	331	22	AAU19218	Human olfactory re
565	5	17.2	309	21	AAU19218	Arabidopsis thalia	638	5	17.2	333	16	AAU19218	Epsilon oploid rec
566	5	17.2	309	21	AAU19218	Arabidopsis thalia	639	5	17.2	333	21	AAU19218	Human G protein-co
567	5	17.2	310	21	AAU19218	Human confluency r	640	5	17.2	333	21	AAU19218	Human mutant G pro
568	5	17.2	310	21	AAU19218	Human confluency r	641	5	17.2	333	22	AAU19218	Novel human diago
569	5	17.2	310	21	AAU19218	Murine confluency	642	5	17.2	333	22	AAU19218	Novel human diago
570	5	17.2	310	21	AAU19218	Human PRO1868 prot	643	5	17.2	333	22	AAU19218	Novel human secret
571	5	17.2	310	21	AAU19218	PRO1868, an A33 an	644	5	17.2	333	22	AAU19218	Human G protein-co
572	5	17.2	310	21	AAU19218	Human IGFAM-6 immu	645	5	17.2	333	22	AAU19218	Human G-protein re
573	5	17.2	310	22	AAU19218	Human polypeptide,	646	5	17.2	333	22	AAU19218	Human G-protein re
574	5	17.2	310	22	AAU19218	Human polypeptide,	647	5	17.2	333	22	AAU19218	Human G-protein re
575	5	17.2	310	22	AAU19218	Human PRO1868 poly	648	5	17.2	334	22	AAU19218	Novel human diago
576	5	17.2	310	22	AAU19218	Human PRO1868 prot	649	5	17.2	336	19	AAU19218	Human secreted pro
577	5	17.2	310	22	AAU19218	Secreted protein e	650	5	17.2	337	20	AAU19218	Human secreted pro
578	5	17.2	310	22	AAU19218	Secreted protein e	651	5	17.2	337	21	AAU19218	Human MATH-2 prote
579	5	17.2	310	22	AAU19218	Secreted protein e	652	5	17.2	337	21	AAU19218	Arabidopsis thalia
580	5	17.2	311	21	AAU19218	Human secreted pro	653	5	17.2	337	21	AAU19218	Arabidopsis thalia
581	5	17.2	311	21	AAU19218	Human secreted pro	654	5	17.2	337	22	AAU19218	Human transcriptio
582	5	17.2	311	21	AAU19218	Human secreted pro	655	5	17.2	337	22	AAU19218	Mouse atonal homol
583	5	17.2	311	22	AAU19218	Propionibacterium	656	5	17.2	337	22	AAU19218	Mouse atonal homol
584	5	17.2	312	20	AAU19218	Porphyrinomas ging	657	5	17.2	337	22	AAU19218	Human atonal prote
585	5	17.2	312	22	ABG24541	Novel human diago	658	5	17.2	338	20	AAU19218	Human ras carboxy-
586	5	17.2	312	22	ABG24541	Novel human diago	659	5	17.2	338	22	AAU19218	Novel human diago
587	5	17.2	314	18	AAU19218	Staphylococcus aur	660	5	17.2	339	15	AAU19218	E.coli OmpA signal
588	5	17.2	314	22	ABG04948	Novel human diago	661	5	17.2	339	22	AAU19218	Novel human diago
589	5	17.2	315	22	AAU19218	Escherichia coli p	662	5	17.2	339	22	AAU19218	Gene #13 associate
590	5	17.2	316	21	AAU19218	Arabidopsis thalia	663	5	17.2	340	21	AAU19218	Zea mays protein f
591	5	17.2	316	22	AAU19218	Human olfactory re	664	5	17.2	341	22	AAU19218	Drosophila melanog
592	5	17.2	316	22	AAU19218	Human olfactory re	665	5	17.2	343	22	AAU19218	Novel human diago
593	5	17.2	316	22	AAU19218	Human olfactory re	666	5	17.2	344	21	AAU19218	Arabidopsis thalia
594	5	17.2	316	22	AAU19218	Human OR-like poly	667	5	17.2	346	22	AAU19218	Novel human diago
595	5	17.2	318	22	ABG01336	Novel human diago	668	5	17.2	346	22	ABG01336	Novel human diago

669	5	17.2	346	22	ABG26933	Novel human diagno	742	5	17.2	383	21	AAG14537	Arabidopsis thalia
670	5	17.2	349	21	AAV69293	A canine growth ho	743	5	17.2	383	21	AAG38967	Arabidopsis thalia
671	5	17.2	350	22	ABG25910	Novel human diagno	744	5	17.2	383	22	AAO13577	Human polypeptide
672	5	17.2	351	21	AAG06147	Arabidopsis thalia	745	5	17.2	385	22	AAO75148	Human colon cancer
673	5	17.2	352	22	ABG26964	Novel human diagno	746	5	17.2	386	22	ABG07122	Novel human diagno
674	5	17.2	353	18	AAW19608	Pig growth hormone	747	5	17.2	386	22	ABG18957	Novel human diagno
675	5	17.2	353	18	AAW19215	Swine growth hormo	748	5	17.2	386	22	ABG24268	Novel human diagno
676	5	17.2	354	21	AAG06997	Arabidopsis thalia	749	5	17.2	386	22	ABG24562	Novel human diagno
677	5	17.2	355	22	ABG07162	Novel human diagno	750	5	17.2	387	21	AAW19413	A straight-chain c
678	5	17.2	355	22	ABG07393	Novel human diagno	751	5	17.2	387	21	AAG21920	Arabidopsis thalia
679	5	17.2	355	22	AAW40136	Human polypeptide	752	5	17.2	387	21	AAG29298	Arabidopsis thalia
680	5	17.2	357	22	ABG02684	Novel human diagno	753	5	17.2	388	21	AAG46587	Arabidopsis thalia
681	5	17.2	357	22	AAW78691	Human protein SEQ	754	5	17.2	388	22	ABG24522	Novel human diagno
682	5	17.2	357	22	AAG82133	S. epidermidis ope	755	5	17.2	390	16	AAW83825	p47(phox) protein.
683	5	17.2	358	12	AAW10614	Pseudomonas glumae	756	5	17.2	391	22	AAW83825	Novel human diagno
684	5	17.2	358	13	AAW20177	P. glumae PGL lipas	757	5	17.2	392	22	ABG26926	Novel human diagno
685	5	17.2	358	13	AAW46834	Arabidopsis thalia	758	5	17.2	392	22	ABG07397	Novel human diagno
686	5	17.2	358	22	ABG26956	Novel human diagno	759	5	17.2	393	21	AAG14536	Drosophila gustato
687	5	17.2	358	22	AAU28020	Novel human diagno	760	5	17.2	393	21	AAG14536	Arabidopsis thalia
688	5	17.2	359	17	AAW88870	Novel human secret	761	5	17.2	393	21	AAG38966	Arabidopsis thalia
689	5	17.2	359	17	AAW88871	Sardinian tomato y	762	5	17.2	394	22	ABG69937	Drosophila melanog
690	5	17.2	359	17	AAW88872	Sardinian tomato y	763	5	17.2	394	22	ABG13232	Novel human diagno
691	5	17.2	359	22	ABG53920	Arabidopsis thalia	764	5	17.2	395	21	AAG06117	Novel human diagno
692	5	17.2	360	12	AAW10438	Drosophila melanog	765	5	17.2	395	21	AAG11791	Arabidopsis thalia
693	5	17.2	361	12	AAW10437	Pseudomonas glumae	766	5	17.2	395	22	ABG21990	Arabidopsis thalia
694	5	17.2	361	18	AAW19217	Pseudomonas glumae	767	5	17.2	395	22	ABG07392	Novel human diagno
695	5	17.2	361	22	ABG22401	Human growth hormo	768	5	17.2	396	19	AAW53868	Gravin polypeptide
696	5	17.2	362	14	AAW14536	Novel human diagno	769	5	17.2	396	21	AAW18408	Human gravin PKA r
697	5	17.2	362	18	AAW19610	Preprokinase 2 tox	770	5	17.2	396	22	ABG07136	Novel human diagno
698	5	17.2	362	21	AAW46584	Human growth hormo	771	5	17.2	396	22	ABG26939	Novel human diagno
699	5	17.2	362	22	ABG16456	Arabidopsis thalia	772	5	17.2	397	22	ABG07395	Novel human diagno
700	5	17.2	363	22	AAU52767	Novel human diagno	773	5	17.2	397	22	ABG21077	Novel human diagno
701	5	17.2	364	8	AAW70310	Propionibacterium	774	5	17.2	398	22	AAW93699	Human protein sequ
702	5	17.2	364	8	AAW70199	Sequence of porcin	775	5	17.2	398	20	AAW07112	Colon cancer assoc
703	5	17.2	364	10	AAW90299	Pseudomonas cepaci	776	5	17.2	400	21	AAG22284	Arabidopsis thalia
704	5	17.2	364	12	AAW10865	Lipase. Pseudomon	777	5	17.2	400	21	AAG39877	Arabidopsis thalia
705	5	17.2	364	12	AAW11425	Heat-resistant lip	778	5	17.2	400	21	AAG46833	Arabidopsis thalia
706	5	17.2	364	14	AAW34633	Lipase B from pseu	779	5	17.2	402	22	ABG13250	Novel human diagno
707	5	17.2	364	14	AAW39396	Pseudomonas cepaci	780	5	17.2	403	22	ABG07128	Novel human diagno
708	5	17.2	364	15	AAW51900	Thermostable lipas	781	5	17.2	403	22	ABG24592	Novel human diagno
709	5	17.2	364	16	AAW7245	P. cepacia lipase.	782	5	17.2	404	22	ABG12109	Novel human diagno
710	5	17.2	364	18	AAW19613	Rat growth hormone	783	5	17.2	404	22	ABG24543	Novel human diagno
711	5	17.2	364	18	AAW19220	Rat growth hormone	784	5	17.2	405	18	AAW20805	Novel human diagno
712	5	17.2	364	21	AAW54565	A mouse growth hor	785	5	17.2	408	21	AAW91059	H. pylori cytoplas
713	5	17.2	364	22	ABG26941	Novel human diagno	786	5	17.2	409	20	AAW97796	Streptomyces nogal
714	5	17.2	364	22	ABG97377	Rat growth hormone	787	5	17.2	409	21	AAG07514	Arabidopsis thalia
715	5	17.2	365	18	AAW12392	Coliurus hirstus m	788	5	17.2	409	21	AAG53783	Arabidopsis thalia
716	5	17.2	365	21	AAW30784	Arabidopsis thalia	789	5	17.2	412	21	AAG49547	Arabidopsis thalia
717	5	17.2	366	21	AAW90632	Human G protein-co	790	5	17.2	414	22	AAW39238	Human polypeptide
718	5	17.2	366	21	AAW90666	Human mutant G pro	791	5	17.2	415	18	AAW30193	Epidermal differen
719	5	17.2	366	21	AAW70345	Human G protein-co	792	5	17.2	415	22	ABG18127	Novel human diagno
720	5	17.2	366	22	ABG07126	Novel human diagno	793	5	17.2	418	21	AAG29524	Arabidopsis thalia
721	5	17.2	366	22	ABG97376	Rat growth hormone	794	5	17.2	418	22	ABG05015	Novel human diagno
722	5	17.2	366	22	ABG26954	Human G-protein co	795	5	17.2	420	22	ABG62543	Drosophila melanog
723	5	17.2	368	22	ABG10027	Novel human diagno	796	5	17.2	421	22	ABG06116	Arabidopsis thalia
724	5	17.2	369	21	AAW43769	Zea mays protein f	797	5	17.2	421	22	ABG25522	Novel human diagno
725	5	17.2	369	22	AAU62756	Propionibacterium	798	5	17.2	422	22	ABG67118	Drosophila melanog
726	5	17.2	369	22	ABG07159	Novel human diagno	799	5	17.2	422	22	ABG21846	Novel human diagno
727	5	17.2	369	22	ABG26954	Novel human diagno	800	5	17.2	423	21	AAW53325	Arabidopsis thalia
728	5	17.2	370	17	AAW98617	Borna disease viru	801	5	17.2	423	22	ABG16849	Novel human diagno
729	5	17.2	370	22	ABG65635	Drosophila melanog	802	5	17.2	425	18	AAW27163	Human TRAF inhibit
730	5	17.2	371	22	ABG07385	Novel human diagno	803	5	17.2	425	21	AAW988237	Human I-TRAF prote
731	5	17.2	371	22	ABG24561	Novel human diagno	804	5	17.2	425	22	ABW70926	Drosophila melanog
732	5	17.2	371	22	ABG26925	Novel human diagno	805	5	17.2	427	22	ABG20627	Novel human diagno
733	5	17.2	373	22	ABG26946	Novel human diagno	806	5	17.2	427	22	AAW41024	Human polypeptide
734	5	17.2	375	22	ABG07389	Novel human diagno	807	5	17.2	428	18	AAW20730	H. pylori cytoplas
735	5	17.2	377	20	AAW34359	Porphyromonas ging	808	5	17.2	428	22	ABG24526	Novel human diagno
736	5	17.2	377	22	ABG24555	Novel human diagno	809	5	17.2	428	22	ABG29008	Novel human diagno
737	5	17.2	377	22	AAW48382	I. ricinus salivar	810	5	17.2	428	22	ABW72624	Human angiotensin
738	5	17.2	378	22	AAO12915	Human polypeptide	811	5	17.2	429	22	ABW67374	Drosophila melanog
739	5	17.2	379	22	ABG21678	Novel human diagno	812	5	17.2	430	22	ABW66369	Drosophila melanog
740	5	17.2	379	22	ABG26952	Novel human diagno	813	5	17.2	437	20	AAW34759	C. pneumoniae 3-de
741	5	17.2	383	21	AAW06124	Drosophila insulin	814	5	17.2	437	22	ABG26966	Novel human diagno

815	5	17.2	437	22	AA069151	Human Acetylcholin	888	5	17.2	463	21	AA06588	Arabidopsis thalia
816	5	17.2	437	22	AA069151	Putative P. abyssi	889	5	17.2	463	21	AA06588	Arabidopsis thalia
817	5	17.2	438	22	AA069151	Novel human diagno	890	5	17.2	463	21	AA06588	Arabidopsis thalia
818	5	17.2	439	22	AA069151	Novel human diagno	891	5	17.2	465	22	AB05796	Drosophila melanog
819	5	17.2	439	22	AA069151	Novel human diagno	892	5	17.2	465	22	AB05796	Novel human diagno
820	5	17.2	439	22	AA069151	Novel human diagno	893	5	17.2	467	12	AA01133	Recombinant chromo
821	5	17.2	439	22	AA069151	Novel human diagno	894	5	17.2	467	12	AA01133	Phytase. Aspergil
822	5	17.2	440	22	AA069151	Drosophila melanog	895	5	17.2	467	15	AA04636	Phytase. Aspergil
823	5	17.2	441	21	AA069151	Aspergillus niger	896	5	17.2	467	15	AA04636	A. niger phytase.
824	5	17.2	441	21	AA069151	Aspergillus niger	897	5	17.2	467	18	AA06500	Aspergillus niger
825	5	17.2	441	21	AA069151	Aspergillus niger	898	5	17.2	467	18	AA06500	Aspergillus niger
826	5	17.2	441	21	AA069151	Aspergillus niger	899	5	17.2	467	20	AA06500	Aspergillus niger
827	5	17.2	441	21	AA069151	Aspergillus niger	900	5	17.2	467	20	AA06500	Consensus phytase-
828	5	17.2	441	21	AA069151	Aspergillus niger	901	5	17.2	467	21	AA06500	A. ficuum phytase
829	5	17.2	442	22	AA069151	Aspergillus niger	902	5	17.2	467	21	AA06500	Consensus phytase
830	5	17.2	442	22	AA069151	Novel human diagno	903	5	17.2	467	21	AA06500	Arabidopsis thalia
831	5	17.2	442	22	AA069151	Novel human diagno	904	5	17.2	467	21	AA06500	Phytase-7, a deriv
832	5	17.2	443	21	AA069151	Arabidopsis thalia	905	5	17.2	467	22	AB05027	Novel human diagno
833	5	17.2	443	21	AA069151	Arabidopsis thalia	906	5	17.2	467	22	AB05027	Consensus phytase-
834	5	17.2	443	22	AA069151	Novel human diagno	907	5	17.2	469	21	AA06500	Arabidopsis thalia
835	5	17.2	443	22	AA069151	Streptomyces galii	908	5	17.2	469	22	AA06500	Human colon cancer
836	5	17.2	444	22	AA069151	Drosophila melanog	909	5	17.2	470	21	AA06500	Arabidopsis thalia
837	5	17.2	444	22	AA069151	Aspergillus niger	910	5	17.2	470	21	AA06500	Arabidopsis thalia
838	5	17.2	445	22	AA069151	Novel human diagno	911	5	17.2	470	22	AA06500	Propionibacterium
839	5	17.2	446	22	AA069151	Novel human diagno	912	5	17.2	471	18	AA06500	H. pylori cytoplas
840	5	17.2	448	22	AA069151	Novel human diagno	913	5	17.2	471	22	AB02976	Novel human diagno
841	5	17.2	449	14	AA069151	Acetylcholine rece	914	5	17.2	471	22	AB02976	Novel human diagno
842	5	17.2	449	15	AA069151	Alpha subunit of T	915	5	17.2	471	22	AB02976	Novel human diagno
843	5	17.2	449	15	AA069151	Acetylcholine rece	916	5	17.2	472	21	AA06500	Arabidopsis thalia
844	5	17.2	449	17	AA069151	Acetylcholine rece	917	5	17.2	472	21	AA06500	Tobacco DI proteas
845	5	17.2	449	22	AA069151	Human hspG13 polyp	918	5	17.2	473	19	AA06500	Novel human diagno
846	5	17.2	449	22	AA069151	Aspergillus niger	919	5	17.2	473	22	AA06500	Novel signal trans
847	5	17.2	449	22	AA069151	Aspergillus niger	920	5	17.2	476	22	AA06500	Novel human diagno
848	5	17.2	453	19	AA069151	Gravin polypeptide	921	5	17.2	477	22	AA06500	Novel human secret
849	5	17.2	453	21	AA069151	Human gravin PKA r	922	5	17.2	481	22	AA06500	Novel human diagno
850	5	17.2	453	22	AA069151	Drosophila melanog	923	5	17.2	481	22	AA06500	Angiotensin conver
851	5	17.2	453	22	AA069151	Drosophila melanog	924	5	17.2	482	22	AB04941	Novel human diagno
852	5	17.2	453	22	AA069151	Novel human diagno	925	5	17.2	482	22	AA06941	Carrot extensin le
853	5	17.2	454	22	AA069151	Drosophila melanog	926	5	17.2	482	22	AA06941	Carrot extensin le
854	5	17.2	454	22	AA069151	ERA binding domain	927	5	17.2	483	21	AA06500	Arabidopsis thalia
855	5	17.2	454	22	AA069151	ERA binding domain	928	5	17.2	483	21	AA06500	Arabidopsis thalia
856	5	17.2	455	11	AA069151	Chicken acetylchol	929	5	17.2	483	22	AB05223	Novel human diagno
857	5	17.2	456	11	AA069151	Human acetylchol	930	5	17.2	483	22	AB05223	Novel human diagno
858	5	17.2	456	11	AA069151	Mouse acetylchol	931	5	17.2	484	21	AA06500	Arabidopsis thalia
859	5	17.2	457	6	AA069151	Calf acetylchol	932	5	17.2	484	21	AA06500	Arabidopsis thalia
860	5	17.2	457	21	AA069151	Human acetyl choli	933	5	17.2	484	22	AB07112	Novel human diagno
861	5	17.2	457	21	AA069151	Human acetylchol	934	5	17.2	484	22	AB07112	Novel human diagno
862	5	17.2	458	22	AA069151	Novel human diagno	935	5	17.2	488	20	AA06500	Rice serine palmit
863	5	17.2	458	22	AA069151	Novel human diagno	936	5	17.2	488	20	AA06500	Rice serine palmit
864	5	17.2	458	22	AA069151	Novel human diagno	937	5	17.2	489	20	AA06500	Protein involved i
865	5	17.2	458	22	AA069151	Novel human diagno	938	5	17.2	489	20	AA06500	Babesia microti BM
866	5	17.2	459	22	AA069151	Novel human diagno	939	5	17.2	491	19	AA06500	Babesia microti an
867	5	17.2	459	22	AA069151	S. epidermidis ope	940	5	17.2	492	20	AA06500	B. microti BMN1-4
868	5	17.2	460	11	AA069151	Fish acetylchol	941	5	17.2	492	21	AA06500	Escherichia coli p
869	5	17.2	460	20	AA069151	Human TIE ligand N	942	5	17.2	492	22	AB05201	Novel human diagno
870	5	17.2	460	21	AA069151	Hydrophobic domain	943	5	17.2	494	22	AB05201	Novel human diagno
871	5	17.2	460	21	AA069151	Human protein PRO1	944	5	17.2	494	22	AB05201	Novel human diagno
872	5	17.2	460	21	AA069151	Human PRO179 prote	945	5	17.2	499	22	AB05201	Drosophila melanog
873	5	17.2	460	21	AA069151	Human TIE ligand F	946	5	17.2	500	12	AA06500	Cl inhibitor mutel
874	5	17.2	460	21	AA069151	Human PRO179 antit	947	5	17.2	500	12	AA06500	Cl inhibitor mutel
875	5	17.2	460	21	AA069151	Human liver angio	948	5	17.2	500	12	AA06500	Cl inhibitor mutel
876	5	17.2	460	21	AA069151	Fragment of human	949	5	17.2	500	12	AA06500	Cl inhibitor mutel
877	5	17.2	460	21	AA069151	Human growth facto	950	5	17.2	500	12	AA06500	Cl inhibitor mutel
878	5	17.2	460	21	AA069151	Scarfase 3 angio	951	5	17.2	500	12	AA06500	Recombinant Cl inh
879	5	17.2	460	21	AA069151	Scarfase 3 angio	952	5	17.2	500	18	AA06500	Recombinant Cl inh
880	5	17.2	460	22	AA069151	Human protein invo	953	5	17.2	500	18	AA06500	Recombinant Cl inh
881	5	17.2	460	22	AA069151	Human immunostimul	954	5	17.2	500	18	AA06500	Recombinant Cl inh
882	5	17.2	460	22	AA069151	Human angio	955	5	17.2	500	18	AA06500	Recombinant Cl inh
883	5	17.2	460	22	AA069151	TSA2306 protein.	956	5	17.2	500	18	AA06500	Recombinant Cl inh
884	5	17.2	461	5	AA069151	Human angiogenesis	957	5	17.2	500	18	AA06500	Recombinant Cl inh
885	5	17.2	461	5	AA069151	Acetyl choline acc	958	5	17.2	500	18	AA06500	Wild-type Cl inh
886	5	17.2	462	22	AA069151	Novel human diagno	959	5	17.2	506	21	AA06500	Arabidopsis thalia
887	5	17.2	462	22	AA069151	Human transport pr	960	5	17.2	506	22	AA06500	Novel human trans

961	5	17.2	509	22	ABG21881	Novel human diagno	1034	5	17.2	550	22	ABG05230	Novel human diagno
962	5	17.2	512	20	AAV20067	B. burgdorferi ant	1035	5	17.2	550	22	ABG29796	Novel human diagno
963	5	17.2	512	22	ABG04308	Novel human diagno	1036	5	17.2	553	22	ABG25033	Novel human diagno
964	5	17.2	512	22	AAE03718	Novel human transf	1037	5	17.2	555	22	ABG24923	Novel human diagno
965	5	17.2	513	21	AAV91276	Group B Streptococ	1038	5	17.2	560	22	AAU40217	Propionibacterium
966	5	17.2	517	22	ABG29821	Novel human diagno	1039	5	17.2	560	22	ABG06797	Novel human diagno
967	5	17.2	519	22	ABG09372	Novel human diagno	1040	5	17.2	560	22	ABG14486	Novel human diagno
968	5	17.2	520	22	AAU41718	Propionibacterium	1041	5	17.2	561	20	AAW82404	Human SRE-ZBP anal
969	5	17.2	521	19	AAW70076	Mycobacterial meth	1042	5	17.2	561	22	AAW93771	Human polypeptide,
970	5	17.2	521	21	AAW57089	Human prostate can	1043	5	17.2	561	22	AAW93949	Human polypeptide,
971	5	17.2	523	21	AAW13324	Arabidopsis thalia	1044	5	17.2	561	22	AAW94525	Human protein sequ
972	5	17.2	523	21	AAW60167	Arabidopsis thalia	1045	5	17.2	563	21	AAW94525	Human protein sequ
973	5	17.2	524	22	ABG04949	Novel human diagno	1046	5	17.2	564	22	ABW62696	Drosophila melanog
974	5	17.2	524	22	AAW41771	Human polypeptide	1047	5	17.2	566	22	AAE03719	Novel human transf
975	5	17.2	525	17	AAW96104	Hepatitis E virus	1048	5	17.2	568	21	AAW59737	Arabidopsis thalia
976	5	17.2	525	21	AAW52594	Eucalyptus grandis	1049	5	17.2	572	22	AAE03717	Novel human transf
977	5	17.2	525	22	AAW92457	Human protein sequ	1050	5	17.2	573	22	AAU08491	Mouse VMGLOM long
978	5	17.2	526	18	AAW30705	Inositol-1-phospha	1051	5	17.2	576	16	AAW71374	Human GRK #2. Hom
979	5	17.2	526	19	AAW79299	Inositol-1-phospha	1052	5	17.2	576	22	AAW84262	Amino acid sequenc
980	5	17.2	526	19	AAW41727	Actinobacillus ple	1053	5	17.2	579	20	AAW25917	Human GPC3 protein
981	5	17.2	526	22	AAW96360	Putative P. abyssi	1054	5	17.2	579	21	AAW95702	Cosmid CHRIM5 enco
982	5	17.2	527	22	ABG00391	Novel human diagno	1055	5	17.2	580	20	AAW25912	Human GPC3 protein
983	5	17.2	529	16	AAW74188	Mouse p78. Mus sp	1056	5	17.2	580	21	AAW35400	Zea mays protein f
984	5	17.2	529	20	AAW20066	B. burgdorferi ant	1057	5	17.2	580	22	ABG25757	Novel human diagno
985	5	17.2	531	22	ABW66934	Drosophila melanog	1058	5	17.2	582	21	AAW44722	Human immune syste
986	5	17.2	531	22	ABG13241	Novel human diagno	1059	5	17.2	582	22	ABG04945	Novel human diagno
987	5	17.2	531	22	AAW25619	Human protein sequ	1060	5	17.2	584	22	ABG04938	Novel human diagno
988	5	17.2	532	19	AAW49700	Human flavin-conta	1061	5	17.2	584	22	ABG05212	Novel human diagno
989	5	17.2	532	21	AAW22283	Arabidopsis thalia	1062	5	17.2	584	22	ABG12694	Novel human diagno
990	5	17.2	532	21	AAW29522	Arabidopsis thalia	1063	5	17.2	586	21	AAW30783	Arabidopsis thalia
991	5	17.2	532	21	AAW39876	Arabidopsis thalia	1064	5	17.2	587	21	AAW49546	Arabidopsis thalia
992	5	17.2	535	22	ABG24530	Novel human diagno	1065	5	17.2	588	21	AAW14794	Arabidopsis thalia
993	5	17.2	536	21	AAW49134	Arabidopsis thalia	1066	5	17.2	588	21	AAW45650	Arabidopsis thalia
994	5	17.2	536	22	ABG26911	Novel human diagno	1067	5	17.2	590	22	ABW58808	Drosophila melanog
995	5	17.2	538	14	AAW1892	Pcp B. Flavobacte	1068	5	17.2	590	22	AAW48181	Thermus thermophil
996	5	17.2	538	17	AAW99485	Pentachlorophenol	1069	5	17.2	590	22	ABG27989	Novel human diagno
997	5	17.2	539	22	ABG02196	Novel human diagno	1070	5	17.2	591	22	ABG13254	Novel human diagno
998	5	17.2	539	22	ABG04950	Novel human diagno	1071	5	17.2	591	22	ABG13426	Novel human diagno
999	5	17.2	540	17	AAW96102	Hepatitis E virus	1072	5	17.2	591	22	ABG15764	Novel human diagno
1000	5	17.2	540	18	AAW19632	Human osteo antivi	1073	5	17.2	591	22	ABG24581	Novel human diagno
1001	5	17.2	540	19	AAW61600	Human extracellula	1074	5	17.2	592	19	AAW56287	Babesia microti BM
1002	5	17.2	540	20	AAW94282	Human extracellula	1075	5	17.2	592	20	AAW24345	Babesia microti an
1003	5	17.2	540	22	AAW68180	Extracellular matr	1076	5	17.2	592	21	AAW30194	B. microti BMNI-8
1004	5	17.2	541	22	ABW60509	Drosophila melanog	1077	5	17.2	594	22	ABG18639	Novel human diagno
1005	5	17.2	541	22	ABG24598	Novel human diagno	1078	5	17.2	594	22	ABG21682	Novel human diagno
1006	5	17.2	543	18	AAW25739	Soybean protoporph	1079	5	17.2	595	22	ABG05214	Novel human diagno
1007	5	17.2	543	20	AAW92498	Glycine max protop	1080	5	17.2	595	22	ABG15755	Novel human diagno
1008	5	17.2	543	21	AAW54100	The chloroplast-ty	1081	5	17.2	595	22	ABG21824	Novel human diagno
1009	5	17.2	543	22	AAW52215	Soybean ppo seq ID	1082	5	17.2	595	22	ABG21874	Novel human diagno
1010	5	17.2	543	22	AAE08753	Soybean protoporph	1083	5	17.2	595	22	ABG29785	Novel human diagno
1011	5	17.2	543	22	AAE08802	Soybean protoporph	1084	5	17.2	597	14	AAW30168	Novel intestinal o
1012	5	17.2	543	22	AAE08803	Soybean protoporph	1085	5	17.2	597	18	AAW20536	H. pylori chapeton
1013	5	17.2	543	22	AAE08804	Soybean protoporph	1086	5	17.2	597	22	ABW60160	Drosophila melanog
1014	5	17.2	543	22	AAE08805	Soybean protoporph	1087	5	17.2	600	21	AAW36724	Arabidopsis thalia
1015	5	17.2	543	22	AAE08806	Soybean protoporph	1088	5	17.2	600	21	AAW49545	Arabidopsis thalia
1016	5	17.2	543	22	AAE08807	Soybean protoporph	1089	5	17.2	601	21	AAW45183	Arabidopsis thalia
1017	5	17.2	543	22	AAE08808	Soybean protoporph	1090	5	17.2	603	21	AAW76839	Netrin protein seq
1018	5	17.2	543	22	AAW93450	Human polypeptide,	1091	5	17.2	603	21	AAW74570	Neisseria gonorrhe
1019	5	17.2	543	22	AAW72906	Soybean protoporph	1092	5	17.2	603	21	AAW74571	Neisseria meningit
1020	5	17.2	543	22	AAE10225	Soybean protoporph	1093	5	17.2	604	21	AAW74572	Human netrin-1. H
1021	5	17.2	543	22	AAE10286	Soybean protoporph	1094	5	17.2	604	18	AAW35946	Human netrin-1. H
1022	5	17.2	543	22	AAE10287	Soybean protoporph	1095	5	17.2	605	16	AAW74186	Chick p78. Gallus
1023	5	17.2	543	22	AAE10288	Soybean protoporph	1096	5	17.2	605	21	AAW35399	Zea mays protein f
1024	5	17.2	543	22	AAE10289	Soybean protoporph	1097	5	17.2	605	21	AAW45620	Arabidopsis thalia
1025	5	17.2	543	22	AAE10290	Soybean protoporph	1098	5	17.2	605	21	AAW76837	Netrin-1 protein s
1026	5	17.2	543	22	AAE10291	Soybean protoporph	1099	5	17.2	605	22	AAW93260	Human protein sequ
1027	5	17.2	543	22	AAE10292	Soybean protoporph	1100	5	17.2	607	21	AAW53324	Arabidopsis thalia
1028	5	17.2	545	14	AAW42742	ced-4 gene product	1101	5	17.2	608	21	AAW59736	Arabidopsis thalia
1029	5	17.2	546	22	AAW15385	Novel human diagno	1102	5	17.2	608	22	ABW64844	Drosophila melanog
1030	5	17.2	546	22	AAW81242	Human AFP protein	1103	5	17.2	612	22	ABG02108	Novel human diagno
1031	5	17.2	547	21	AAW39875	Arabidopsis thalia	1104	5	17.2	613	22	ABW62232	Drosophila melanog
1032	5	17.2	548	21	AAW22282	Arabidopsis thalia	1105	5	17.2	614	21	AAW40231	Arabidopsis thalia
1033	5	17.2	549	17	AAW96092	Hepatitis E virus	1106	5	17.2	616	21	AAW23861	Haemophilus influe

1107	5	17.2	621	21	AAG34266	Arabidopsis thalia
1108	5	17.2	621	22	ABG24599	Novel human diagno
1109	5	17.2	621	22	ABG29119	Novel human diagno
1110	5	17.2	625	22	ABG02187	Novel human diagno
1111	5	17.2	627	22	ABG26953	Novel human diagno
1112	5	17.2	628	22	ABG05449	Novel human diagno
1113	5	17.2	628	22	ABG24554	Novel human diagno
1114	5	17.2	629	22	ABG13446	Novel human diagno
1115	5	17.2	630	22	ABG28655	Novel human diagno
1116	5	17.2	631	22	ABG21704	Novel human diagno
1117	5	17.2	631	22	ABG22173	Novel human diagno
1118	5	17.2	632	16	ABG71373	Human GRK #1. Hom
1119	5	17.2	633	22	ABG04038	Novel human diagno
1120	5	17.2	633	22	ABG13465	Novel human diagno
1121	5	17.2	633	22	ABG22232	C glutamicum prote
1122	5	17.2	634	22	AA359985	Human polypeptide
1123	5	17.2	636	22	AAU02972	Angiotensin conver
1124	5	17.2	636	22	AA94352	Human protein sequ
1125	5	17.2	637	20	AAV25004	Melon MEERS protei
1126	5	17.2	638	18	AAW21012	H. pylori chaperon
1127	5	17.2	638	22	ABG07166	Novel human diagno
1128	5	17.2	643	22	ABG13451	Novel human diagno
1129	5	17.2	643	22	ABG15840	Novel human diagno
1130	5	17.2	643	22	ABG21687	Novel human diagno
1131	5	17.2	643	22	ABG24564	Novel human diagno
1132	5	17.2	644	22	ABG18134	Novel human diagno
1133	5	17.2	645	22	ABG61545	Drosophila melanog
1134	5	17.2	647	21	AAV53000	Human secreted pro
1135	5	17.2	649	22	ABG62131	Drosophila melanog
1136	5	17.2	649	22	ABG61236	Mature human TANGO
1137	5	17.2	651	19	AAW53875	Gravin polypeptide
1138	5	17.2	651	21	AAW18410	Human gravin PKA r
1139	5	17.2	652	22	ABG04021	Novel human diagno
1140	5	17.2	652	22	ABG25758	Novel human diagno
1141	5	17.2	652	22	ABG28386	Novel human diagno
1142	5	17.2	652	22	AAU07895	Human hspG13 polyp
1143	5	17.2	653	22	ABG05219	Novel human diagno
1144	5	17.2	653	22	ABG29788	Novel human diagno
1145	5	17.2	657	22	ABG13201	Novel human diagno
1146	5	17.2	657	22	ABG26912	Novel human diagno
1147	5	17.2	657	22	ABG28395	Novel human diagno
1148	5	17.2	659	14	AAW38787	HEV ORF2 protein.
1149	5	17.2	659	14	AAW39308	Mexico strain HEV
1150	5	17.2	659	18	AAW35827	Hepatitis E virus
1151	5	17.2	659	20	AAW93387	Human HEV ORF 2 pr
1152	5	17.2	659	21	AAW52517	Helicobacter pylor
1153	5	17.2	659	22	AAW33061	C glutamicum prote
1154	5	17.2	660	17	AAW96090	Hepatitis E virus
1155	5	17.2	661	18	AAW15565	Human intracellular
1156	5	17.2	664	17	AAW04329	Human HP-8 antigen
1157	5	17.2	664	22	AAW78753	Human protein SEQ
1158	5	17.2	666	18	AAW15564	Human intracellular
1159	5	17.2	667	22	ABG16901	Novel human diagno
1160	5	17.2	667	22	AAU52908	Novel human secret
1161	5	17.2	668	22	ABG13500	Novel human diagno
1162	5	17.2	671	22	ABG08247	Novel human diagno
1163	5	17.2	671	22	AAW61234	Human TANGO 332 pr
1164	5	17.2	674	22	AAE00416	Tomato hydrolase,
1165	5	17.2	675	22	ABG14106	Novel human diagno
1166	5	17.2	676	21	AAW7469	Fucus vanadium bro
1167	5	17.2	682	22	AAW62036	P. furiosus mcm po
1168	5	17.2	684	21	AAW58217	Lung cancer associ
1169	5	17.2	687	21	AAG34265	Arabidopsis thalia
1170	5	17.2	687	22	ABG14091	Novel human diagno
1171	5	17.2	693	22	ABG28399	Novel human diagno
1172	5	17.2	696	22	ABG61188	Drosophila melanog
1173	5	17.2	696	22	AAU04874	Micromonospora eve
1174	5	17.2	697	22	AAW93191	Human protein sequ
1175	5	17.2	699	14	AAW32549	Tomato hsp80 encod
1176	5	17.2	699	21	AAG43720	Arabidopsis thalia
1177	5	17.2	699	21	AAG47033	Arabidopsis thalia
1178	5	17.2	699	21	AAG47039	Arabidopsis thalia
1179	5	17.2	699	22	ABW71022	Drosophila melanog
1180	5	17.2	700	21	AAG39766	Arabidopsis thalia
1181	5	17.2	702	22	AAU49114	Propionibacterium
1182	5	17.2	702	22	ABG18349	Novel human diagno
1183	5	17.2	704	21	AAW77930	A. thaliana enviro
1184	5	17.2	705	21	AAG39765	Arabidopsis thalia
1185	5	17.2	706	18	AAW22231	S. pombe origin of
1186	5	17.2	707	18	AAW14137	Schizosaccharomyce
1187	5	17.2	710	22	ABG23625	Novel human diagno
1188	5	17.2	714	20	AAW39934	DS9 protein sequen
1189	5	17.2	714	21	AAW10372	Tobacco DS9 protei
1190	5	17.2	715	22	AAW79737	Human protein SEQ
1191	5	17.2	716	22	ABG13499	Novel human diagno
1192	5	17.2	716	22	ABG25880	Novel human diagno
1193	5	17.2	716	22	AAU12219	Human PRO4304 poly
1194	5	17.2	716	22	AAG90527	C glutamicum prote
1195	5	17.2	718	16	AAW85865	WD-40 domain-contg
1196	5	17.2	719	15	AAW51481	Drosophila E(spl)m
1197	5	17.2	719	15	AAW52955	Drosophila E(spl)m
1198	5	17.2	719	22	ABW63831	Drosophila melanog
1199	5	17.2	719	22	ABW66968	Drosophila melanog
1200	5	17.2	723	22	ABG28392	Novel human diagno

ALIGNMENTS

RESULT 1

ABW03187 ID AAB03187 standard; peptide: 21 AA.

XX AAB03187;

XX 23-OCT-2000 (first entry)

XX Human selenoprotein P C-terminal fragment, SEQ ID NO:5.

XX Selenoprotein P fragment; human; programmed cell death inhibitor;
 XX apoptosis inhibitor; antiapoptotic; acquired immunodeficiency syndrome;
 XX AIDS; Parkinson's disease; Alzheimer's disease; myocardial infarct;
 XX cerebral infarct; stroke; reperfusion injury; organ transplant rejection;
 XX arteriosclerosis; redox disorder; immune disorder; cardiovascular;
 XX cerebroprotective; culture medium additive.

XX Homo sapiens.

XX WO2000031131-Al.

XX 02-JUN-2000.

XX 12-NOV-1999; 99WO-JP06322.

XX 19-NOV-1998; 98JP-0347863.

XX (KAGA) CHEMAO-SERO-THERAPEUTIC RES INST.

XX Hirashima M, Maeda H, Nozaki C;

XX WPI; 2000-451677/39.

XX Peptide fragments with cell death inhibitory activity, useful in
 XX preventing and treating apoptosis-associated diseases particularly
 XX caused by stress e.g. Parkinson's disease, Alzheimer's and
 XX arteriosclerosis

XX Example 3; Page 19; 56pp; Japanese.

XX The invention relates to fragments (AAB03183, AAB03184) derived from the
 XX C-terminal 103 amino acids of human selenoprotein P (AAB03188) which act
 XX as inhibitors of apoptosis (programmed cell death), their use in
 XX therapeutic compositions and cell culture media, and antibodies against
 XX them. The invention also relates to a method for studying the phenomenon
 XX of sudden cell death in a human megakaryotic series gemmule cell culture
 XX system involving use of the anti-apoptotic peptides of the invention.

CC Selenoprotein P is a plasma protein which confers protection against
 CC peroxynitrite-mediated oxidation and nitration. The peptide fragments
 CC can be used as prophylactic and therapeutic agents for
 CC apoptosis-associated diseases, including AIDS (acquired immunodeficiency
 CC syndrome), Parkinson's disease, Alzheimer's disease, arteriosclerosis,
 CC myocardial infarct, cerebral infarct, organ transplant rejection, and
 CC reperfusion injury. They may also be used in the prevention or treatment
 CC of disorders relating to redox reactions, or of disorders of the immune
 CC system. They may additionally be used as an additive for cell culture and
 CC for screening cell death activity in vitro. The present sequence
 CC represents a C-terminal fragment of human selenoprotein P used in an
 CC exemplification of the invention.

XX
 SQ Sequence 21 AA;

Query Match 72.4%; Score 21; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.4e-15;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRCINQLLKLPTDSELAPRS 21
 |||||
 Db 1 krcinqllklptdselaprs 21

RESULT 2
 AAB03183
 ID AAB03183 standard; peptide; 29 AA.

AC AAB03183;

DT 23-OCT-2000 (first entry)

XX Human selenoprotein P C-terminal fragment, SEQ ID NO:1.

XX Selenoprotein P fragment; human; programmed cell death inhibitor;
 KW apoptosis inhibitor; antiapoptotic; acquired immunodeficiency syndrome;
 KW AIDS; Parkinson's disease; Alzheimer's disease; myocardial infarct;
 KW cerebral infarct; stroke; reperfusion injury; organ transplant rejection;
 KW arteriosclerosis; redox disorder; immune disorder; cardiovascular;
 KW cerebroprotective; culture medium additive.

OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT Modified-site 22 /note= "Selenocysteine"

XX WO200031131-A1.

XX 02-JUN-2000.

XX 12-NOV-1999; 99WO-JP06322.

XX 19-NOV-1998; 98JP-0347863.

XX (KAGA) CHEMO-SERO-THERAPEUTIC RES INST.

XX Hirashima M, Maeda H, Nozaki C;

XX WPI; 2000-451677/39.

XX Peptide fragments with cell death inhibitory activity, useful in
 PT preventing and treating apoptosis-associated diseases particularly
 PT caused by stress e.g. Parkinson's disease, Alzheimer's and
 PT arteriosclerosis

XX Claim 2; Page 31; 56pp; Japanese.

XX This sequence represents a fragment of human selenoprotein P which
 CC has anti-apoptotic activity. The invention relates to fragments
 CC (AAB03183, AAB03184) derived from the C-terminal 103 amino acids of
 CC human selenoprotein P (AAB03188) which act as inhibitors of apoptosis

CC (programmed cell death), their use in therapeutic compositions and cell
 CC culture media, and antibodies against them. The invention also relates to
 CC a method for studying the phenomenon of sudden cell death in a human
 CC megakaryotic series gemmule cell culture system involving use of the
 CC anti-apoptotic peptides of the invention. Selenoprotein P is a plasma
 CC protein which confers protection against peroxynitrite-mediated
 CC oxidation and nitration. The peptide fragments can be used as
 CC prophylactic and therapeutic agents for apoptosis-associated diseases,
 CC including AIDS (acquired immunodeficiency syndrome), Parkinson's disease,
 CC Alzheimer's disease, arteriosclerosis, myocardial infarct, cerebral
 CC infarct, organ transplant rejection, and reperfusion injury. They may
 CC also be used in the prevention or treatment of disorders relating to
 CC redox reactions, or of disorders of the immune system. They may
 CC additionally be used as an additive for cell culture and for screening
 CC cell death activity in vitro.

SQ Sequence 29 AA;

Query Match 72.4%; Score 21; DB 21; Length 29;
 Best Local Similarity 100.0%; Pred. No. 3.1e-15;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRCINQLLKLPTDSELAPRS 21
 |||||
 Db 1 krcinqllklptdselaprs 21

RESULT 3

AAB03185
 ID AAB03185 standard; peptide; 103 AA.

XX AAB03185;

DT 23-OCT-2000 (first entry)

XX Human selenoprotein P C-terminal 103 residues.

XX Selenoprotein P fragment; human; programmed cell death inhibitor;
 KW apoptosis inhibitor; antiapoptotic; acquired immunodeficiency syndrome;
 KW AIDS; Parkinson's disease; Alzheimer's disease; myocardial infarct;
 KW cerebral infarct; stroke; reperfusion injury; organ transplant rejection;
 KW arteriosclerosis; redox disorder; immune disorder; cardiovascular;
 KW cerebroprotective; culture medium additive.

OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT Modified-site 22 /note= "Selenocysteine"
 FT Modified-site 40 /note= "Selenocysteine"
 FT Modified-site 52 /note= "Selenocysteine"
 FT Modified-site 67 /note= "Selenocysteine"
 FT Modified-site 74 /note= "Selenocysteine"
 FT Modified-site 89 /note= "Selenocysteine"
 FT Modified-site 91 /note= "Selenocysteine"
 FT Modified-site 98 /note= "Selenocysteine"
 FT Modified-site 100 /note= "Selenocysteine"
 FT /note= "Selenocysteine"

XX WO200031131-A1.

XX 02-JUN-2000.

XX 12-NOV-1999; 99WO-JP06322.

XX


```
XX PS Claim 11; Page 2116-2117; 2338pp; English.
XX CC AAF15566 to AAF16505 encode the human prostate cancer associated
XX CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
XX CC The prostate cancer antigens can have neuroprotective, cytostatic,
XX CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
XX CC nephrotropic, antinfecive, gynaecological and antibacterial activities,
XX CC and can be used in gene therapy. The prostate cancer antigen
XX CC polynucleotides may be used for detection of prostate cancer, chromosome
XX CC identification, as chromosome markers, and for numerous other diagnostic
XX CC or research purposes. The prostate cancer antigens may be used to treat
XX CC disorders such as neural, immune, muscular, reproductive,
XX CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
XX CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
XX CC AAB57303 represent sequences used in the exemplification of the present
XX CC invention.
XX PS Sequence 240 AA;

Query Match 72.4%; Score 21; DB 21; Length 240;
Best Local Similarity 100.0%; Pred. No. 1.6e-14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRCINOLLCKLPTDSELAPRS 21
Db 220 krcinqlcklptdselaprs 240
|||||

RESULT 6
AAB03188
ID AAB03188 standard; peptide; 381 AA.
XX AC AAB03188;
XX DT 23-OCT-2000 (first entry)
XX DE Human selenoprotein P.
XX KW Selenoprotein P; human; programmed cell death inhibitor;
XX KW apoptosis inhibitor; antiapoptotic; acquired immunodeficiency syndrome;
XX KW AIDS; Parkinson's disease; Alzheimer's disease; myocardial infarct;
XX KW cerebral infarct; stroke; reperfusion injury; organ transplant rejection;
XX KW arteriosclerosis; redox disorder; immune disorder; cardiovascular;
XX KW cerebroprotective; culture medium additive.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Peptide 1..19
XX FT Protein 19..381
XX FT /note= "Signal peptide"
XX FT Modified-site 281
XX FT /note= "Mature human selenoprotein P"
XX FT Modified-site 299
XX FT /note= "Selenocysteine"
XX FT Modified-site 311
XX FT /note= "Selenocysteine"
XX FT Modified-site 326
XX FT /note= "Selenocysteine"
XX FT Modified-site 333
XX FT /note= "Selenocysteine"
XX FT Modified-site 348
XX FT /note= "Selenocysteine"
XX FT Modified-site 350
XX FT /note= "Selenocysteine"
XX FT Modified-site 357
XX FT /note= "Selenocysteine"
XX FT Modified-site 359
XX FT /note= "Selenocysteine"
XX FT Modified-site 359
XX FT /note= "Selenocysteine"

PN WO200031131-A1.
XX PD 02-JUN-2000.
XX PF 12-NOV-1999; 99WO-JP06322.
XX PR 19-NOV-1998; 98JP-0347863.
XX PA (KAGA ) CHEMO-SERO-THERAPEUTIC RES INST.
XX PI Hirashima M, Maeda H, Nozaki C;
XX WIPI: 2000-451677/39.
XX PT Peptide fragments with cell death inhibitory activity, useful in
XX PT preventing and treating apoptosis-associated diseases particularly
XX PT caused by stress e.g. Parkinson's disease, Alzheimer's and
XX PT arteriosclerosis
XX PS Example 3; Page 20; 56pp; Japanese.
XX CC The invention relates to fragments (AAB03183, AAB03184) derived from the
XX CC C-terminal 103 amino acids of human selenoprotein P (AAB03188) which act
XX CC as inhibitors of apoptosis (programmed cell death), their use in
XX CC therapeutic compositions and cell culture media, and antibodies against
XX CC them. The invention also relates to a method for studying the phenomenon
XX CC of sudden cell death in a human megakaryotic series gammale cell culture
XX CC system involving use of the anti-apoptotic peptides of the invention.
XX CC Selenoprotein P is a plasma protein which confers protection against
XX CC peroxynitrite-mediated oxidation and nitration. The peptide fragments
XX CC can be used as prophylactic and therapeutic agents for
XX CC apoptosis-associated diseases, including AIDS (acquired immunodeficiency
XX CC syndrome), Parkinson's disease, Alzheimer's disease, arteriosclerosis,
XX CC myocardial infarct, cerebral infarct, organ transplant rejection, and
XX CC reperfusion injury. They may also be used in the prevention or treatment
XX CC of disorders relating to redox reactions, or of disorders of the immune
XX CC system. They may additionally be used as an additive for cell culture and
XX CC for screening cell death activity in vitro. The present sequence
XX CC represents human selenoprotein P.
XX PS Sequence 381 AA;
XX SQ

Query Match 72.4%; Score 21; DB 21; Length 381;
Best Local Similarity 100.0%; Pred. No. 2.3e-14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRCINOLLCKLPTDSELAPRS 21
Db 279 krcinqlcklptdselaprs 299
|||||

RESULT 7
AAB03186
ID AAB03186 standard; peptide; 20 AA.
XX AC AAB03186;
XX DT 23-OCT-2000 (first entry)
XX DE Human selenoprotein P C-terminal fragment, SEQ ID NO:4.
XX KW Selenoprotein P fragment; human; programmed cell death inhibitor;
XX KW apoptosis inhibitor; antiapoptotic; acquired immunodeficiency syndrome;
XX KW AIDS; Parkinson's disease; Alzheimer's disease; myocardial infarct;
XX KW cerebral infarct; stroke; reperfusion injury; organ transplant rejection;
XX KW arteriosclerosis; redox disorder; immune disorder; cardiovascular;
XX KW cerebroprotective; culture medium additive.
XX OS Homo sapiens.
XX PN WO200031131-A1.
XX XX
```

PD 02-JUN-2000.
XX
PF 12-NOV-1999; 99WO-JP06322.
XX
PR 19-NOV-1998; 98JP-0347863.
XX
PA (KAGA) CHEMO-SERO-THERAPEUTIC RES INST.
XX
PI Hirashima M, Maeda H, Nozaki C;
XX WPI; 2000-451677/39.
DR
XX
XX Peptide fragments with cell death inhibitory activity, useful in
PT preventing and treating apoptosis-associated diseases particularly
PT caused by stress e.g. Parkinson's disease, Alzheimer's and
PT arteriosclerosis
XX
PS Example 4; Page 11; 56pp; Japanese.
XX
CC The invention relates to fragments (AAB03183, AAB03184) derived from the
CC C-terminal 103 amino acids of human selenoprotein P (AAB03188) which act
CC as inhibitors of apoptosis (programmed cell death), their use in
CC therapeutic compositions and cell culture media, and antibodies against
CC them. The invention also relates to a method for studying the phenomenon
CC of sudden cell death in a human megakaryotic series gemmule cell culture
CC system involving use of the anti-apoptotic peptides of the invention.
CC Selenoprotein P is a plasma protein which confers protection against
CC peroxynitrite-mediated oxidation and nitration. The peptide fragments
CC can be used as prophylactic and therapeutic agents for
CC apoptosis-associated diseases, including AIDS (acquired immunodeficiency
CC syndrome), Parkinson's disease, Alzheimer's disease, arteriosclerosis,
CC myocardial infarct, cerebral infarct, organ transplant rejection, and
CC reperfusion injury. They may also be used in the prevention or treatment
CC of disorders relating to redox reactions, or of disorders of the immune
CC system. They may additionally be used as an additive for cell culture and
CC for screening cell death activity in vitro. The present sequence
CC represents a C-terminal fragment of human selenoprotein P used in an
CC exemplification of the invention.
XX
SQ Sequence 20 AA;

Query Match 69.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.6e-14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRCINQLICKLPTDSELAPR 20
|||
Db 1 krcinqllicklptdselapr 20

RESULT 8
AAU17709
ID AAU17709 standard; Protein; 51 AA.
XX
AC AAU17709;
XX
DT 07-NOV-2001 (first entry)
XX
DE Novel human respiratory antigen #25.
XX
KW Human; respiratory antigen; respiratory disorder; throat disorder;
KW lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
KW anti allergic; anti asthmatic; anti inflammatory; olfactory;
KW respiratory active.
XX
OS Homo sapiens.
XX
PN WO200155448-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01333.

XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.

PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-476224/51.
 DR N-PSDB; AAS27893.
 DR

PT Isolated polypeptide for treating, preventing and/ or prognosing

PT disorders related to the respiratory system including respiratory
 PT cancers and also for testing and detection e.g. diagnosis -
 XX Claim 11; SED ID No 327; 546pp; English.
 PS
 XX The present invention relates to the isolation of novel human
 CC respiratory antigens, and cDNA (AAS27869-AAS28159) and genomic
 CC sequences encoding for these polypeptides. The sequences of the
 CC invention are useful for preventing, treating and/or prognosing
 CC disorders related to the respiratory system including throat
 CC disorders (e.g. vocal cord paralysis, tonsillitis, and laryngitis),
 CC lung disorders e.g. pneumonia, allergic disorders e.g. asthma,
 CC pleurisy, cystic fibrosis, emphysema, nose disorders and cancers of
 CC the respiratory tissues e.g. lung cancer. The polynucleotide sequences
 CC of the invention are useful in gene therapy and antisense therapy.
 CC AAU17685-AAU17975 represent novel human respiratory antigens.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 51 AA;

 Query Match 20.7%; Score 6; DB 22; Length 51;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 15 SELAPR 20
 Db 37 selapr 42

 RESULT 9
 AAG33566
 ID AAG33566 standard; Protein; 106 AA.
 XX
 AC AAG33566;
 XX
 DT 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 40692.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 23-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.


```
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 20.7%; Score 6; DB 21; Length 112;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDSELA 18
Db 73 tdsela 78

RESULT 11
AAG33564
ID AAG33564 standard; Protein; 116 AA.
XX
AC AAG33564;
XX
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 40690.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
```

PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.

PR 29-OCT-1999; 99US-0162142.

Query Match 20.7%; Score 6; DB 21; Length 116;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDSELA 18
| | | | |
Db 77 tdsele 82

RESULT 12
ABG27736
ID ABG27736 standard; Protein; 120 AA.

XX AC ABG27736;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #27727.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.

XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI: 2001-639362/73.
XX DR N-PSDB; AAS91923.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -

XX PS Claim 20; SEQ ID No 58095; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human
XX CC diagnostic amino acid sequences of the invention.

XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 120 AA;

Query Match 20.7%; Score 6; DB 22; Length 120;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LACKLP 12
| | | | |
Db 21 llacklp 26

RESULT 13
ABG27735
ID ABG27735 standard; Protein; 134 AA.

XX AC ABG27735;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #27726.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.

XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI: 2001-639362/73.
XX DR N-PSDB; AAS91922.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -

XX PS Claim 20; SEQ ID No 58094; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human
XX CC diagnostic amino acid sequences of the invention.

XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 134 AA;

Query Match	20.7%;	Score 6;	DB 22;	Length 134;	
Best Local Similarity	100.0%;	Pred. No. 58;			
Matches 6;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	7	LLCKLP 12			
Db	37	licklp 42			
RESULT 14					
AAAG29772					
ID	AAAG29772	standard; Protein; 170 AA.			
XX	AC	AAAG29772;			
XX	DT	17-OCT-2000 (first entry)			
XX	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 35479.			
XX	KW	Protein identification; signal transduction pathway; metabolic pathway;			
KW	hybridisation assay; genetic mapping; gene expression control; promoter;				
KW	termination sequence.				
XX	OS	Arabidopsis thaliana.			
XX	XX				
PN	EP1033405-A2.				
XX	PD				
XX	06-SEP-2000.				
PF	25-FEB-2000; 2000EP-0301439.				
XX	25-FEB-1999;	99US-0121825.			
PR	05-MAR-1999;	99US-0123180.			
PR	09-MAR-1999;	99US-0123548.			
PR	23-MAR-1999;	99US-0125788.			
PR	25-MAR-1999;	99US-0126264.			
PR	29-MAR-1999;	99US-0126785.			
PR	01-APR-1999;	99US-0127462.			
PR	06-APR-1999;	99US-0128234.			
PR	08-APR-1999;	99US-0128714.			
PR	16-APR-1999;	99US-0129845.			
PR	19-APR-1999;	99US-0130077.			
PR	21-APR-1999;	99US-0130449.			
PR	23-APR-1999;	99US-0130510.			
PR	23-APR-1999;	99US-0130891.			
PR	28-APR-1999;	99US-0131449.			
PR	30-APR-1999;	99US-0132048.			
PR	30-APR-1999;	99US-0132407.			
PR	04-MAY-1999;	99US-0132484.			
PR	05-MAY-1999;	99US-0132485.			
PR	06-MAY-1999;	99US-0132486.			
PR	07-MAY-1999;	99US-0132487.			
PR	11-MAY-1999;	99US-0132863.			
PR	11-MAY-1999;	99US-0134256.			
PR	14-MAY-1999;	99US-0134218.			
PR	14-MAY-1999;	99US-0134219.			
PR	14-MAY-1999;	99US-0134221.			
PR	14-MAY-1999;	99US-0134370.			
PR	18-MAY-1999;	99US-0134768.			
PR	19-MAY-1999;	99US-0134941.			
PR	20-MAY-1999;	99US-0135124.			
PR	21-MAY-1999;	99US-0135353.			
PR	24-MAY-1999;	99US-0135629.			
PR	25-MAY-1999;	99US-0136021.			
PR	27-MAY-1999;	99US-0136392.			
PR	28-MAY-1999;	99US-0136782.			
PR	01-JUN-1999;	99US-0137222.			
PR	03-JUN-1999;	99US-0137528.			
PR	04-JUN-1999;	99US-0137502.			
PR	07-JUN-1999;	99US-0137724.			
PR	08-JUN-1999;	99US-0138094.			
PR	10-JUN-1999;	99US-0138540.			
PR	10-JUN-1999;	99US-0138847.			
PR	14-JUN-1999;	99US-0139119.			
PR	16-JUN-1999;	99US-0139452.			
PR	16-JUN-1999;	99US-0139453.			
PR	17-JUN-1999;	99US-0139492.			
PR	18-JUN-1999;	99US-0139454.			
PR	18-JUN-1999;	99US-0139455.			
PR	18-JUN-1999;	99US-0139456.			
PR	18-JUN-1999;	99US-0139457.			
PR	18-JUN-1999;	99US-0139458.			
PR	18-JUN-1999;	99US-0139459.			
PR	18-JUN-1999;	99US-0139460.			
PR	18-JUN-1999;	99US-0139461.			
PR	18-JUN-1999;	99US-0139462.			
PR	18-JUN-1999;	99US-0139463.			
PR	18-JUN-1999;	99US-0139750.			
PR	18-JUN-1999;	99US-0139763.			
PR	21-JUN-1999;	99US-0139817.			
PR	22-JUN-1999;	99US-0139899.			
PR	23-JUN-1999;	99US-0140353.			
PR	23-JUN-1999;	99US-0140354.			
PR	24-JUN-1999;	99US-0140695.			
PR	28-JUN-1999;	99US-0140823.			
PR	29-JUN-1999;	99US-0140991.			
PR	30-JUN-1999;	99US-0141287.			
PR	01-JUL-1999;	99US-0141842.			
PR	01-JUL-1999;	99US-0142154.			
PR	02-JUL-1999;	99US-0142055.			
PR	06-JUL-1999;	99US-0142390.			
PR	08-JUL-1999;	99US-0142803.			
PR	09-JUL-1999;	99US-0142920.			
PR	12-JUL-1999;	99US-0142977.			
PR	13-JUL-1999;	99US-0143542.			
PR	14-JUL-1999;	99US-0143624.			
PR	15-JUL-1999;	99US-0144005.			
PR	16-JUL-1999;	99US-0144085.			
PR	16-JUL-1999;	99US-0144086.			
PR	19-JUL-1999;	99US-0144331.			
PR	19-JUL-1999;	99US-0144332.			
PR	19-JUL-1999;	99US-0144333.			
PR	19-JUL-1999;	99US-0144334.			
PR	19-JUL-1999;	99US-0144335.			
PR	20-JUL-1999;	99US-0144352.			
PR	20-JUL-1999;	99US-0144632.			
PR	21-JUL-1999;	99US-0144884.			
PR	21-JUL-1999;	99US-0144814.			
PR	21-JUL-1999;	99US-0145086.			
PR	21-JUL-1999;	99US-0145088.			
PR	22-JUL-1999;	99US-0145085.			
PR	22-JUL-1999;	99US-0145087.			
PR	22-JUL-1999;	99US-0145089.			
PR	22-JUL-1999;	99US-0145192.			
PR	23-JUL-1999;	99US-0145145.			
PR	23-JUL-1999;	99US-0145218.			
PR	23-JUL-1999;	99US-0145244.			
PR	26-JUL-1999;	99US-0145276.			
PR	27-JUL-1999;	99US-0145913.			
PR	27-JUL-1999;	99US-0145918.			
PR	27-JUL-1999;	99US-0145919.			
PR	28-JUL-1999;	99US-0145951.			
PR	02-AUG-1999;	99US-0146386.			
PR	02-AUG-1999;	99US-0146388.			
PR	02-AUG-1999;	99US-0146389.			
PR	03-AUG-1999;	99US-0147038.			
PR	04-AUG-1999;	99US-0147204.			
PR	04-AUG-1999;	99US-0147302.			
PR	05-AUG-1999;	99US-0147192.			
PR	05-AUG-1999;	99US-0147260.			
PR	06-AUG-1999;	99US-0147303.			
PR	06-AUG-1999;	99US-0147416.			
PR	09-AUG-1999;	99US-0147493.			
PR	09-AUG-1999;	99US-0147935.			
PR	10-AUG-1999;	99US-0148171.			

PR 11-AUG-1999;	99US-0148319.	PR 25-FEB-1999;	99US-0121825.
PR 12-AUG-1999;	99US-0148341.	PR 05-MAR-1999;	99US-0123180.
PR 13-AUG-1999;	99US-0148565.	PR 09-MAR-1999;	99US-0123548.
PR 13-AUG-1999;	99US-0148684.	PR 23-MAR-1999;	99US-0125788.
PR 16-AUG-1999;	99US-0149368.	PR 25-MAR-1999;	99US-0126264.
PR 17-AUG-1999;	99US-0149175.	PR 29-MAR-1999;	99US-0126785.
PR 18-AUG-1999;	99US-0149426.	PR 01-APR-1999;	99US-0127462.
PR 20-AUG-1999;	99US-0149722.	PR 08-APR-1999;	99US-0128234.
PR 20-AUG-1999;	99US-0149723.	PR 16-APR-1999;	99US-0128714.
PR 20-AUG-1999;	99US-0149729.	PR 19-APR-1999;	99US-0129845.
PR 23-AUG-1999;	99US-0149902.	PR 21-APR-1999;	99US-0130077.
PR 23-AUG-1999;	99US-0149930.	PR 23-APR-1999;	99US-0130449.
PR 25-AUG-1999;	99US-0150566.	PR 28-APR-1999;	99US-0130510.
PR 26-AUG-1999;	99US-0150884.	PR 30-APR-1999;	99US-0130891.
PR 27-AUG-1999;	99US-0151065.	PR 05-MAY-1999;	99US-0131449.
PR 27-AUG-1999;	99US-0151066.	PR 06-MAY-1999;	99US-0132485.
PR 27-AUG-1999;	99US-0151080.	PR 07-MAY-1999;	99US-0132486.
PR 30-AUG-1999;	99US-0151303.	PR 11-MAY-1999;	99US-0132487.
PR 31-AUG-1999;	99US-0151438.	PR 14-MAY-1999;	99US-0132863.
PR 01-SEP-1999;	99US-0151930.	PR 14-MAY-1999;	99US-0134218.
PR 07-SEP-1999;	99US-0152363.	PR 14-MAY-1999;	99US-0134221.
PR 10-SEP-1999;	99US-0153070.	PR 18-MAY-1999;	99US-0134768.
PR 13-SEP-1999;	99US-0153758.	PR 19-MAY-1999;	99US-0134941.
PR 15-SEP-1999;	99US-0154018.	PR 20-MAY-1999;	99US-0135124.
PR 16-SEP-1999;	99US-0154039.	PR 21-MAY-1999;	99US-0135353.
PR 20-SEP-1999;	99US-0154779.	PR 25-MAY-1999;	99US-0136021.
PR 22-SEP-1999;	99US-0155139.	PR 27-MAY-1999;	99US-0136392.
PR 23-SEP-1999;	99US-0155486.	PR 28-MAY-1999;	99US-0136782.
PR 24-SEP-1999;	99US-0155659.	PR 01-JUN-1999;	99US-0137222.
PR 28-SEP-1999;	99US-0156458.	PR 03-JUN-1999;	99US-0137528.
PR 29-SEP-1999;	99US-0156596.	PR 04-JUN-1999;	99US-0137502.
PR 04-OCT-1999;	99US-0157117.	PR 07-JUN-1999;	99US-0137724.
PR 05-OCT-1999;	99US-0157753.	PR 08-JUN-1999;	99US-0138094.
PR 06-OCT-1999;	99US-0157865.	PR 10-JUN-1999;	99US-0138540.
PR 07-OCT-1999;	99US-0158029.	PR 14-JUN-1999;	99US-0138847.
PR 12-OCT-1999;	99US-0158232.	PR 16-JUN-1999;	99US-0139119.
PR 12-OCT-1999;	99US-0158369.	PR 16-JUN-1999;	99US-0139452.
PR 13-OCT-1999;	99US-0159293.	PR 17-JUN-1999;	99US-0139453.
PR 13-OCT-1999;	99US-0159294.	PR 18-JUN-1999;	99US-0139454.
PR 14-OCT-1999;	99US-0159295.	PR 18-JUN-1999;	99US-0139455.
PR 14-OCT-1999;	99US-0159329.	PR 18-JUN-1999;	99US-0139456.
PR 14-OCT-1999;	99US-0159330.		
PR 14-OCT-1999;	99US-0159331.		
PR 14-OCT-1999;	99US-0159637.		
PR 14-OCT-1999;	99US-0159638.		
PR 18-OCT-1999;	99US-0159584.		
PR 21-OCT-1999;	99US-0160741.		
PR 21-OCT-1999;	99US-0160767.		
PR 21-OCT-1999;	99US-0160768.		
PR 21-OCT-1999;	99US-0160770.		
PR 21-OCT-1999;	99US-0160814.		
PR 21-OCT-1999;	99US-0160815.		
PR 22-OCT-1999;	99US-0160980.		
PR 22-OCT-1999;	99US-0160981.		
PR 22-OCT-1999;	99US-0160989.		
PR 25-OCT-1999;	99US-0161404.		
PR 25-OCT-1999;	99US-0161405.		
PR 25-OCT-1999;	99US-0161406.		
PR 26-OCT-1999;	99US-0161359.		
PR 26-OCT-1999;	99US-0161360.		
PR 26-OCT-1999;	99US-0161361.		
PR 28-OCT-1999;	99US-0161920.		
PR 28-OCT-1999;	99US-0161992.		
PR 28-OCT-1999;	99US-0161993.		
PR 29-OCT-1999;	99US-0162142.		

Query Match
Best Local Similarity 100.0%; DB 21; Length 170;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDSELA 18
|||||
Db 67 tdsela 72

RESULT 15

AAG29771

ID AAG29771 standard; Protein; 176 AA.

XX AAG29771;

XX AAG29771;

DT 17-OCT-2000 (first entry)

XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 35478.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.

PR

PR 05-MAR-1999; 99US-0123180.

PR

PR 09-MAR-1999; 99US-0123548.

PR

PR 23-MAR-1999; 99US-0125788.

PR

PR 25-MAR-1999; 99US-0126264.

PR

PR 29-MAR-1999; 99US-0126785.

PR

PR 01-APR-1999; 99US-0127462.

PR

PR 08-APR-1999; 99US-0128234.

PR

PR 16-APR-1999; 99US-0128714.

PR

PR 19-APR-1999; 99US-0129845.

PR

PR 21-APR-1999; 99US-0130077.

PR

PR 23-APR-1999; 99US-0130449.

PR

PR 28-APR-1999; 99US-0130510.

PR

PR 30-APR-1999; 99US-0130891.

PR

PR 05-MAY-1999; 99US-0131449.

PR

PR 06-MAY-1999; 99US-0132485.

PR

PR 07-MAY-1999; 99US-0132486.

PR

PR 11-MAY-1999; 99US-0132487.

PR

PR 14-MAY-1999; 99US-0132863.

PR

PR 14-MAY-1999; 99US-0134218.

PR

PR 14-MAY-1999; 99US-0134221.

PR

PR 18-MAY-1999; 99US-0134768.

PR

PR 19-MAY-1999; 99US-0134941.

PR

PR 20-MAY-1999; 99US-0135124.

PR

PR 21-MAY-1999; 99US-0135353.

PR

PR 25-MAY-1999; 99US-0135629.

PR

PR 27-MAY-1999; 99US-0136021.

PR

PR 28-MAY-1999; 99US-0136392.

PR

PR 01-JUN-1999; 99US-0136782.

PR

PR 03-JUN-1999; 99US-0137222.

PR

PR 04-JUN-1999; 99US-0137502.

PR

PR 07-JUN-1999; 99US-0137724.

PR

PR 08-JUN-1999; 99US-0138094.

PR

PR 10-JUN-1999; 99US-0138540.

PR

PR 14-JUN-1999; 99US-0138847.

PR

PR 16-JUN-1999; 99US-0139119.

PR

PR 16-JUN-1999; 99US-0139452.

PR

PR 17-JUN-1999; 99US-0139453.

PR

PR 18-JUN-1999; 99US-0139454.

PR

PR 18-JUN-1999; 99US-0139455.

PR

PR 18-JUN-1999; 99US-0139456.

PR

PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 23-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 20.7%; Score 6; DB 21; Length 176;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDESLA 18
| | | | |
Db 73 tdsela 78

RESULT 16
AAG29770
ID AAG29770 standard; Protein; 180 AA.
XX
AC AAG29770;


```
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157293.
PR 06-OCT-1999; 99US-0157553.
PR 07-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 20.7%; Score 6; DB 21; Length 180;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDSELA 18
Db 77 tdsla 82

RESULT 17
AAU02498
ID AAU02498 standard; Protein; 276 AA.
XX
AC AAU02498;
XX
DT 07-SEP-2001 (first entry)
DE Murine trinucleotide repeat protein (TRP).
XX Trinucleotide repeat protein; TRP; T243; embryonic stem cell; ES;
KW transgenic animal; knockout mouse; triplet repeat expansion;
```

```
KW fragile X syndrome; Huntington's disease; mouse; murine.
XX Mus sp.
XX WO200130798-A1.
XX 03-MAY-2001.
XX 26-OCT-2000; 2000WO-US29382.
XX 26-OCT-1999; 99US-0161488.
XX (DELT-) DELTAGEN INC.
XX Klein R, Matthews W, Moore M, Allen KD;
XX WPI; 2001-300473/31.
XX N-PSDB; AAS05278.
XX Novel transgenic animals useful as animal model for characterization of
XX function of a gene encoding trinucleotide repeat proteins (TRPs),
XX contains heterozygous disruption in a gene encoding TRP -
XX Claim 62; Fig 12; 106pp; English.
XX The present sequence represents murine trinucleotide repeat protein
XX (TRP). The invention describes methods of producing embryonic stem (ES)
XX cells comprising a heterozygous disruption in a target DNA sequence
XX (preferably gene T243) encoding a TRP and of producing a knockout mouse
XX comprising a homozygous disruption in a gene encoding TRP, where the
XX disruption inhibits the production of the wild type TRP. The invention
XX also relates to identifying agents capable of affecting a phenotype of
XX a knockout mouse. Also described are methods of determining whether
XX expansion of the trinucleotide repeat in a gene encoding TRP produces a
XX phenotypic change. The transgenic animals and the cells are useful for
XX identifying compounds capable of ameliorating disease symptoms, and as
XX test substrates for the identification of drugs, pharmaceuticals,
XX therapies and interventions which may be effective in treating
XX trinucleotide repeat disorders e.g. fragile X syndrome and Huntington's
XX disease. The animal models for trinucleotide repeat disorders are ideal
XX model systems to study the progression of disease in vivo, the molecular
XX basis of these diseases and show the features observed in human disease.
XX Using the mice, it is possible to model both the pathogenic mechanism and
XX the trinucleotide repeat instability in the mouse.
XX Sequence 276 AA;

Query Match 20.7%; Score 6; DB 22; Length 276;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SELAPR 20
Db 5 selapr 10

RESULT 18
AAG55022
ID AAG55022 standard; Protein; 311 AA.
XX
AC AAG55022;
XX
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 70445.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
```


PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 20.7%; Score 6; DB 21; Length 311;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDSELA 18
 |||||
 Db 208 tdsela 213

RESULT 19
 AAB19382
 ID AAB19382 standard; Protein; 334 AA.

XX AC AAB19382;

XX DT 06-MAR-2001 (first entry)

XX DE Amino acid sequence of a human lipoxigenase protein.

XX KW Human; lipoxigenase; leukotriene; lipid; chemotactic agent;
 KW inflammation; smooth muscle contraction; asthma; eye disease; arthritis;
 KW lung disease; cancer; acne; psoriasis.

XX OS Homo sapiens.

XX PN WO200061765-A2.

XX PD 19-OCT-2000.

XX PF 12-APR-2000; 2000WO-US09657.

XX PR 12-APR-1999; 99US-0128817.

PR 24-AUG-1999; 99US-0150454.
 XX (LEXI-) LEXICON GENETICS INC.

XX PI Turner CA, Zambrowicz B, Nehls M, Friedrich G, Sands AT;

XX DR WPI; 2000-665134/64.

DR N-PSDB; AAC61750.

XX Novel polynucleotides encoding human lipoxigenase proteins useful for
 PT producing transgenic animals preferably mouse -

XX PS Claim 3; Page 64-65; 83pp; English.

XX CC AAB19379-92 represent novel human lipoxigenase proteins. Lipoxigenases
 CC oxidise lipids to produce leukotrienes. Leukotrienes bind cognate
 CC receptors and trigger biological effects. Leukotrienes influence a
 CC variety of biological processes, and can serve as, inter alia, potent
 CC chemotactic agents and mediators of inflammation, smooth muscle
 CC contractions, etc.. Lipoxigenases and leukotrienes are implicated in
 CC a variety of diseases and disorders, such as asthma, eye diseases,
 CC arthritis, lung disease, cancer, acne, psoriasis, etc...

XX Sequence 334 AA;

Query Match 20.7%; Score 6; DB 21; Length 334;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 LPTDSE 16

|||||

Db 225 lptdse 230

RESULT 20

AAY06628

ID AAY06628 standard; Protein; 359 AA.

XX AC AAY06628;

XX DT 26-OCT-1999 (first entry)

XX DE C. elegans tyrosylprotein sulfotransferase TPST-B.

XX KW TPST-B; tyrosylprotein sulfotransferase; tyrosine O-sulfation;
 KW post-translation modification.

XX OS Caenorhabditis elegans.

XX PN WO9938980-A2.

XX PD 05-AUG-1999.

XX PF 28-JAN-1999; 99WO-US01756.

XX PR 09-SEP-1998; 98US-0150141.

XX PR 29-JAN-1998; 98US-0072994.

XX PR 09-SEP-1998; 98US-0150133.

XX PA (OKLA) UNIV OKLAHOMA STATE.

XX PI Moore KL;

XX DR WPI; 1999-494094/41.

XX DR N-PSDB; AAX87691.

XX PT Human, murine and Caenorhabditis elegans tyrosylprotein
 PT sulfotransferases, useful for post-translational tyrosine sulfation

XX PS Claim 4; Page 122; 123pp; English.

XX CC This sequence represents tyrosylprotein sulfotransferase TPST-B of

CC Caenorhabditis elegans. TPST enzymes catalyze the transfer of
 CC sulfate from 3'-phosphoadenosine 5'-phosphosulfate to tyrosine
 CC residues within highly acidic motifs of polypeptides. C. elegans
 CC TPST-B has type II transmembrane topology. It shows 39% identity
 CC and 62% similarity to 54% C. elegans TPST-A. Dual isotopes of TPST
 CC have been identified in human, mouse and C. elegans (see AAY06623-28).
 CC TPST polypeptides can be produced by recombinant methods and used
 CC for post-translational tyrosine sulfation of proteins or peptides.
 CC Tyrosine O-sulfation plays a role in protein-protein interactions
 CC in several biological systems.

XX
 SQ Sequence 359 AA;

Query Match 20.7%; Score 6; DB 20; Length 359;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SELAPR 20
 |||||
 Db 161 selapr 166

RESULT 21
 AAY84309
 ID AAY84309 standard; Protein; 359 AA.
 XX
 AC AAY84309;
 XX
 DT 12-JUL-2000 (first entry)
 XX
 DE A tyrosylprotein sulfotransferase (TPST) polypeptide.
 XX
 KW Tyrosylprotein sulfotransferase; TPST; tyrosine O-sulfation;
 KW P-selectin glycoprotein ligand-1; PSGL-1; anti-inflammatory;
 KW neutrophil binding.
 XX
 OS Caenorhabditis elegans.
 XX
 PN WO200014250-A1.
 XX
 PD 16-MAR-2000.
 XX
 PF 23-JUL-1999; 99WO-US16750.
 XX
 PR 09-SEP-1998; 98US-0150133.
 PR 09-SEP-1998; 98US-0150141.
 PR 28-JAN-1999; 99WO-US01756.
 XX
 PA (OKLA) UNIV OKLAHOMA STATE.
 XX
 PI Moore KL;
 XX
 DR WPI; 2000-271057/23.
 DR N-PSDB; AAZ99702.
 XX
 PT New human or murine tyrosylprotein sulfotransferase, used for in vitro
 PT or in vivo sulfation of proteins and to screen for anti-inflammatory
 PT agents -
 XX
 PS Claim 37; Page 123-124; 141pp; English.
 XX
 CC The present sequence represents a tyrosylprotein sulfotransferase (TPST)
 CC polypeptide. TPST proteins catalyze the transfer of sulfate from
 CC 3'-phosphoadenosine-5'-phosphosulfate to tyrosine residues (tyrosine
 CC O-sulfation) in proteins. TPST polypeptides are used for in vitro or
 CC in vivo sulfation of proteins or peptides, particularly P-selectin
 CC glycoprotein ligand-1 (PSGL-1) or its functional fragments. They are
 CC also used to raise specific antibodies, and are useful as immunoassay
 CC reagents, and to identify specific inhibitors. These inhibitors are
 CC potentially useful as anti-inflammatory agents, by inhibiting PSGL-1
 CC mediated binding of neutrophils to endothelial cells. The TPST
 CC polynucleotide is useful for identifying polymorphisms that may be

CC associated with a disease phenotype.
 XX
 SQ Sequence 359 AA;

Query Match 20.7%; Score 6; DB 21; Length 359;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SELAPR 20
 |||||
 Db 161 selapr 166

RESULT 22
 AAB32501
 ID AAB32501 standard; Protein; 404 AA.
 XX
 AC AAB32501;
 XX
 DT 19-JAN-2001 (first entry)
 XX
 DE S. lavendulae Mit G encoded protein sequence.
 XX
 KW Mitomycin; biosynthesis; mitosome ring system; antibiotic; anti-cancer;
 KW anti-inflammatory; immune-enhancer; immunosuppressant; asthma;
 KW chronic obstructive pulmonary disease; respiratory inflammation;
 KW fungicide; pesticide.
 XX
 OS Streptomyces lavendulae.
 XX
 PN WO200053737-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 10-MAR-2000; 2000WO-US06394.
 XX
 PR 12-MAR-1999; 99US-0266965.
 XX
 PA (MINU) UNIV MINNESOTA.
 PA (SHER/) SHERMAN D H.
 PA (MAOY/) MAO Y.
 PA (VARO/) VAROGLU M.
 PA (HEMM/) HE M.
 PA (SHEL/) SHERDON P C.
 XX
 PI Sherman DH, Mao Y, Varoglu M, He M, Sheldon PC;
 XX
 DR WPI; 2000-601980/57.
 DR N-PSDB; AAC55808.
 XX
 PT Novel nucleic acid molecule comprising mitomycin biosynthetic gene
 PT cluster useful for cloning mitomycin biosynthetic genes for elucidating
 PT the molecular basis of mitosome ring system biosynthesis -
 XX
 PS Disclosure; Page 345-346; 399pp; English.
 XX
 CC This invention relates to isolated and purified nucleic acid molecules
 CC from the mitomycin biosynthetic gene cluster. Mitomycins are a group of
 CC natural products that contain a variety of functional groups, including
 CC amino benzoquinone and axiridine ring systems. The S. lavendulae
 CC mitomycin biosynthetic gene cluster comprises 47 mitomycin genes
 CC spanning 55kb of DNA. The invention includes an expression cassette
 CC comprising a mitomycin biosynthetic gene operably linked to a promoter,
 CC and host cells transformed with the cassette. The nucleotide, and protein
 CC sequences and the transformed host cells of the invention result in
 CC antiasthmatic, antiinflammatory, cytostatic, immunomodulatory, and
 CC antibiotic activities. The nucleotide sequences are used to elucidate the
 CC molecular basis for the biosynthesis of the mitosome ring system, as well
 CC as to engineer the biosynthesis of novel natural products, e.g.
 CC antibiotics, anti-inflammatory agents, anti-cancer agents,
 CC immune-enhancers, immunosuppressants, agents to treat asthma, chronic
 CC obstructive pulmonary disease as well as other disease involving

CC respiratory inflammation, or cholesterol-lowering agents or as crop
CC protection agents (e.g. fungicides or insecticides) as well as
CC biopolymers, e.g., in packaging or biomedical applications, or to engineer
CC PHA monomer syntheses. Sequences AAC55782-C55881, AAC55815-C55849 and
CC AAB32485-B32542 represent mitomycin biosynthetic gene cluster DNA
CC sequences and encoded proteins. Sequences AAC55812-C55814,
CC AAC55850-C55856 and AAC55862-C55869 represent PCR primers used in the
CC cloning of the mitomycin biosynthetic genes.
XX
SQ Sequence 404 AA;

Query Match 20.7%; Score 6; DB 21; Length 404;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELAPRS 21

|||||

Db 67 elaprs 72

RESULT 23

AAG21644

ID AAG21644 standard; Protein; 423 AA.

XX AC AAG21644;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 24274.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.

```
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157855.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.

PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 20.7%; Score 6; DB 21; Length 423;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 12 PTDSEL 17
Db 4 ptdsel 9
|||||

RESULT 24
AAU19832
ID AAU19832 standard; Protein; 427 AA.
XX AC AAU19832;
XX DT 06-DEC-2001 (first entry)
XX DE Human novel extracellular matrix protein, Seq ID No 482.
XX KW Human; secreted extracellular matrix protein; immunomodulatory;
KW Anti-HIV; antianemic; antirheumatic; antisclerotic; cardiant; vascular;
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;
KW antiaizheimers; immune/autoimmune disease; HIV infection; anaemia;
KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;
KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;
KW Sezary syndrome; Gaucher's disease; neurological diseases;
KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
KW cardiac arrest; tachycardia; angina; infection; corneal infections;
KW wound healing; immunogen; gene therapy; antisense; food additive.
XX OS Homo sapiens.
XX PN WO200155368-A1.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01348.
XX PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
```

PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 06-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.

PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-465572/50.
XX N-PSDB; AAS31403.
XX
XX Nucleic acid molecules encoding human secreted extracellular matrix
PT proteins, used in preventing, treating or ameliorating a disorder, e.g.
PT Alzheimer's and Parkinson's diseases and cancers -
XX
PS Claim 11; SEQ ID No 482; 577pp; English.
XX
XX The invention relates to isolated nucleic acid molecules encoding
CC novel human secreted extracellular matrix proteins (SPs). The
CC polynucleotides and proteins are used to prevent, treat a medical
CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
CC chickens or sheep. For example, disorders associated with decreased
CC expression of SPs. The SP polynucleotide or a vector expressing them may
CC be administered to treat diseases by gene therapy. Antisense molecules
CC may be administered to down regulate expression of SPs by binding with
CC the cells own genes and preventing their expression. The polynucleotides
CC may also be used as DNA probes in diagnostic assays. The SPs may also be
CC used as antigens to produce antibodies and to identify modulators
CC (agonists and antagonists) of the SPs. The anti-(SP) antibodies and
CC antagonists may also be used to down regulate expression and activity of
CC SP and as diagnostic agents for detecting the presence of SPs in samples.
CC The disorders include for example: immune/autoimmune diseases (e.g. HIV
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
CC Parkinson's disease) cardio-/cerebrovascular disorders (e.g. cardiac
CC arrest, tachycardia and angina), infections caused by bacteria, viruses
CC and fungi and ocular disorders (e.g. corneal infections). Other uses
CC include wound healing, maintenance of organs before transplantation,

Query Match 20.7%; Score 6; DB 22; Length 427;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 ELAPRS 21
|||||
Db 185 elaprs 190

RESULT 25
AAB19384
ID AAB19384 standard; Protein; 460 AA.
XX
AC AAB19384;
XX
XX 06-MAR-2001 (first entry)
DT
XX
XX Amino acid sequence of a human lipoxxygenase protein.
DE
KW Human; lipoxxygenase; leukotriene; lipid; chemotactic agent;
KW Inflammation; smooth muscle contraction; asthma; eye disease; arthritis;
KW lung disease; cancer; acne; psoriasis.
XX
OS Homo sapiens.
XX
XX WO200061765-A2.
PN
XX
XX 19-OCT-2000.
PD
XX
XX 12-APR-2000; 2000WO-US09657.
PF
XX
XX 12-APR-1999; 99US-0128817.
PR
XX 24-AUG-1999; 99US-0150454.
PR
XX (LEXI-) LEXICON GENETICS INC.
PA
XX
XX Turner CA, Zambrowicz B, Nehls M, Friedrich G, Sands AT;
PI
XX
XX WPI; 2000-665134/64.
XX
DR N-PSDB; AAC61752.
XX
XX Novel polynucleotides encoding human lipoxxygenase proteins useful for
PT producing transgenic animals preferably mouse -
PT
XX
XX Claim 3; Page 68-69; 83pp; English.
PS
XX AAB19379-92 represent novel human lipoxxygenase proteins. Lipoxxygenases
CC oxidise lipids to produce leukotrienes. Leukotrienes bind cognate
CC receptors and trigger biological effects. Leukotrienes influence a
CC variety of biological processes, and can serve as, inter alia, potent
CC chemotactic agents and mediators of inflammation, smooth muscle
CC contractions, etc.. Lipoxxygenases and leukotrienes are implicated in
CC a variety of diseases and disorders, such as asthma, eye diseases,
CC arthritis, lung disease, cancer, acne, psoriasis, etc..
XX
SQ Sequence 460 AA;

Query Match 20.7%; Score 6; DB 21; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 LPTDSE 16
|||||
Db 225 lptdse 230

RESULT 26
AAR46065
ID AAR46065 standard; protein; 478 AA.
XX
XX AAR46065;
AC
XX
XX 18-JUL-1994 (first entry)
DT
XX

Query Match 20.7%; Score 6; DB 15; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 PTDSEL 17
|||||
Db 353 ptdsel 358

RESULT 27
AAR79025
ID AAR79025 standard; protein; 478 AA.
XX
XX AAR79025;
AC
XX
XX 22-MAR-1996 (first entry)
DT
XX
XX Mature taka-amylase A.
DE
XX
XX Wild type; neopullulanase; B. stearothermophilus; mutant;
KW food industry; modification; hydrophobicity; replacement
KW insertion; deletion.
XX
OS Aspergillus oryzae.
XX
XX Key Location/Qualifiers
FH Disulfide-bond 30...38
FT Disulfide-bond 150...164
FT Disulfide-bond 240...283
FT Disulfide-bond 439...474

DE
XX Mutant alpha-amylase.
KW Methionine substitution; stability; activity; detergent;
KW dishwashing agents; liquifaction agents.
XX
XX Aspergillus oryzae.
OS
XX WO9402597-A.
PN
XX 03-FEB-1994.
PD
XX
XX 06-JUL-1993; 93WO-DK00230.
PF
XX
XX 23-JUL-1992; 92DK-0000946.
PR
XX 16-DEC-1992; 92DK-0001503.
PR
XX 15-MAR-1993; 93DK-0000292.
PR
XX (NOVO) NOVO-NORDISK AS.
PA
XX Bisgard-frantzen H, Svendsen A;
PI
XX WPI; 1994-048855/06.
DR
XX
XX Mutant alpha-amylase from Bacillus species comprising a
PT methionine substitution - with improved stability and activity at
PT low pH, for use in detergents, dishwashing agents and
PT liquifaction agents
PT
XX
XX Claim 1; Page 7; 20pp; English.
PS
XX
XX The sequence os that of the Asoerigillus oryzae alpha amylase, sold
CC commercially as FUNGAMYL (TM) by Novo Nordisk A/S. The sequence can
CC be mutated by substitution of one or more of its methionine residues
CC for any amino acid other than cysteine. The mutant alpha-amylase
CC exhibits a better activity level and better stability in the
CC presence of oxidising agents than previous mutant alpha amylases,
CC and improved thermostability at moderately low pH. The enzyme can
CC be used as an additive for detergents, dishwashing agents and
CC liquifaction agents.
XX
XX Sequence 478 AA;

XX JP07177891-A.
 XX 18-JUL-1995.
 PD 31-OCT-1994; 94JP-0288658.
 XX 12-NOV-1993; 93JP-0306096.
 XX (EZAK) EZAKI GLICO CO.
 PA (NIDE) NEC CORP.
 XX WPI; 1995-279919/37.
 DR
 XX Modifying a transferase by enhancing hydrophobicity of a selected
 PT site - increases transfer activity, also new mutant
 PT neo-pullulanase(s)
 XX
 PS Disclosure; Page 10-11; 18pp; English.
 XX
 CC This sequence represents the mature form of taka-amylase from A. oryzae.
 CC This sequence was used in a method for the generation of mutant
 CC pullulanases for use in the food industry (see also AAR79026-28). The
 CC wild type pullulanase enzyme was modified by the method of the
 CC invention for enhancing the hydrophobicity of a selected site of the
 CC pullulanase. The method comprises replacement of a group in the
 CC selected site with a hydrophobic group, replacement of an amino acid
 CC with a hydrophobic amino acid, and/or insertion or deletion of a
 CC hydrophobic amino acid from the selected site. The method was used
 CC to produce neopullulanases Y377F, S422V and M375L.
 XX
 SQ Sequence 478 AA;
 Query Match 20.7%; Score 6; DB 16; Length 478;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 12 PTDSEL 17
 Db 353 ptdsel 358
 |||||
 RESULT 28
 AAR72450
 ID AAR72450 standard; Protein; 478 AA.
 XX
 AC AAR72450;
 XX
 DT 01-DEC-1995 (first entry)
 XX
 DE Aspergillus oryzae alpha amylase (mature protein).
 XX
 KW Alpha amylase; variant; enzyme; detergent; additive; dishwashing;
 KW washing; Bacillus licheniformis; Bacillus amyloliquefaciens;
 KW Bacillus stearothermophilus; dyeing; bleaching; scouring; textile;
 KW thermostable.
 XX
 OS Aspergillus oryzae.
 XX
 PN WO9510603-A.
 XX
 PD 20-APR-1995.
 XX
 PF 05-OCT-1994; 94WO-DK00370.
 XX
 PR 08-OCT-1993; 93DK-0001133.
 PR 02-FEB-1994; 94DK-0000140.
 XX
 PA (NOVO) NOVO-NORDISK AS.
 XX
 PI Svendsen A, Thellersen M;
 PI Van der zee P, Bisgardfrantzen H, Borchert T;

XX WPI; 1995-161790/21.
 XX
 DR New Bacillus derived alpha-amylase variants - having amino acid
 XX modifications to improve washing and/or dishwashing performance
 PT
 PT
 XX Disclosure; Page 75-76; 105pp; English.
 PS
 XX Variant alpha amylase enzymes which have improved washing and/or
 CC as detergent additives. The enzymes have one or more amino acid
 CC residues added, deleted or substituted. The variants can also be
 CC used for textile desizing prior to scouring, bleaching and dyeing.
 CC The variants have improved thermostability, acid/alkaline stability;
 CC low temperature optimum; pH optimum; higher hydrolysis velocity and
 CC improved tolerance to other composition constituents, e.g. oxidation
 CC agents.
 XX
 SQ Sequence 478 AA;
 Query Match 20.7%; Score 6; DB 16; Length 478;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 12 PTDSEL 17
 Db 353 ptdsel 358
 |||||
 RESULT 29
 AAR78270
 ID AAR78270 standard; Protein; 478 AA.
 XX
 AC AAR78270;
 XX
 DT 17-JAN-1996 (first entry)
 XX
 DE Aspergillus oryzae alpha amylase (mature protein).
 XX
 KW Alpha amylase; oxidation; desizing; bleaching; scouring; fabric;
 KW starch; thermostable; methionine; Bacillus licheniformis;
 KW Bacillus amyloliquefaciens; Bacillus stearothermophilus;
 KW Aspergillus oryzae.
 XX
 OS Aspergillus oryzae.
 XX
 PN WO9521247-A1.
 XX
 PD 10-AUG-1995.
 XX
 PF 05-OCT-1994; 94WO-DK00371.
 XX
 PR 02-FEB-1994; 94DK-0000141.
 XX
 PA (NOVO) NOVO-NORDISK AS.
 XX
 PI Marcher D, Nilsson TE, Pedersen HH, Toft AH;
 XX
 DR WPI; 1995-283767/37.
 XX
 PT Use of an oxidation stable alpha-amylase - for simultaneous desizing
 PT and bleaching or scouring of fabrics contg. starch or starch derivs.
 XX
 PS Disclosure; Page 25-26; 37pp; English.
 XX
 CC Oxidation stable alpha amylases can be used for the simultaneous
 CC desizing and bleaching or scouring of a fabric comprising starch or
 CC starch derivatives. They exhibit a better heat stability,
 CC especially in the presence of oxidising agents. They are obtained
 CC from a parent alpha amylase by replacing one or more methionine
 CC residues with any amino acid different from Cys or Met, preferably
 CC Leu, Thr, Ala, Gly, Ser, Ile or Asp. The parent alpha amylase is
 CC pref. derived from a Bacillus species, although alpha amylases of

CC fungal origin can also be used. This sequence is the wild type
CC (unmodified) alpha amylase of *Aspergillus oryzae*.

SQ Sequence 478 AA;

Query Match 20.7%; Score 6; DB 16; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 PTDSEL 17
|||||

Db 353 ptdsel 358

RESULT 30

AAW14500
ID AAW14500 standard; protein; 478 AA.

XX AC AAW14500;

XX DT 04-JUN-1997 (first entry)

XX DE *Aspergillus oryzae* alpha-amylase (mature protein).

XX KW alpha-amylase; mature; *Bacillus licheniformis*; termamyl; fungamyl;
KW *Aspergillus oryzae*; *Bacillus amyloliquefaciens*; altered property;
KW calcium dependency; substrate binding; stability; pH optimum;
KW thermostability; cleavage; oligosaccharide substrate; dishwashing;
KW washing; detergent additive; fabric desizing; starch liquefaction;
KW sweetener; ethanol production; variant.

XX OS *Aspergillus oryzae*.

XX FH Key Location/Qualifiers

FT Misc-difference 13..45

FT /label= loop 1 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 7-23 of AAW14499 is deleted or
FT replaced with a fragment corresponding to
FT this fragment; claim 33"

FT Misc-difference 14..40

FT /label= loop 1 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 8-18
FT of AAW14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 35" Misc-difference 28..42

FT /label= loop 1 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 12-19 of AAW14499 is deleted or
FT replaced with a fragment corresponding to
FT this fragment; claim 30"

FT Misc-difference 32..38

FT /label= loop 1 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 14-15
FT of AAW14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 32" Misc-difference 66..84

FT /label= loop 2 modification region

FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 44-57 of AAW14499 is deleted or
FT replaced with a fragment corresponding to
FT this fragment; claim 18"

FT Misc-difference 70..78

FT /label= loop 2 modification region
FT /note= "preferred region where at least one amino acid

FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 48-51
FT of AAW14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 20" Misc-difference 98..210

FT /label= loop 3 modification region

FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 117-185 of AAW14499 is deleted
FT or replaced with a fragment corresponding to
FT this fragment; claim 24"

FT Misc-difference 102..206

FT /label= loop 3 modification region

FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 121-181
FT of AAW14499 is deleted or replaced with a fragment
FT corresponding to this fragment; claim
FT 26" Misc-difference 121..181

FT /note= "at least one amino acid residue of a parent

FT alpha-amylase (used as a template for a variant)
FT corresponding to this fragment is deleted or
FT replaced with a fragment corresponding to
FT 102-206 of AAW14499; claim 41"

FT Misc-difference 121..174

FT /note= "preferred region where at least one amino acid

FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to this
FT fragment is deleted or replaced with a fragment
FT corresponding to 102-199 of AAW14499; claim 42"

FT Misc-difference 165..177

FT /label= loop 3 modification region

FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 195-202 of AAW14499 is deleted
FT or replaced with a fragment corresponding to
FT this fragment; claim 21"

FT Misc-difference 166..173

FT /label= loop 3 modification region

FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 196-198
FT of AAW14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 23" Misc-difference 181..184

FT /note= "an amino acid fragment corresponding to this

FT region is deleted from the parent sequence of a
FT variant Fungamyl; claim 43"

FT Misc-difference 291..313

FT /label= loop 8 modification region

FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 322-346 of AAW14498 is deleted
FT or replaced with a fragment corresponding to
FT this fragment; claim 36"

FT Misc-difference 297..313

FT /label= loop 8 modification region

FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 325-345
FT of AAW14498 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 38"

XX WO9623874-A1.

XX 08-AUG-1996.

XX 05-FEB-1996; 96WO-DK00057.

XX 10-NOV-1995; 95DK-0001256.

XX 03-FEB-1995; 95DK-0001128.

XX 23-OCT-1995; 95DK-0001128.

XX (NOVO) NOVO-NORDISK AS.
 XX Bisgard-frantzen H, Borchert TV, Svendsen A;
 XX WPI; 1996-371424/37.
 XX Alpha-amylase variants and methods of production - have altered
 PT properties such as calcium dependency, substrate binding and
 PT stability
 XX
 XX Disclosure; Page 87-88; 171pp; English.
 XX
 XX The present sequence is the mature Aspergillus oryzae alpha-amylase (A).
 CC Variants of parent termamyl- and fungamyl-like alpha-amylases (and
 CC methods of constructing them) are claimed. Examples of variants are
 CC featured above. The variants have altered properties such as calcium
 CC dependency, substrate binding and stability. Also one or more proline or
 CC cysteine residues in the variant is modified or replaced with a
 CC non-proline or non-cysteine residue such as alanine. The variants can be
 CC used for (dish)washing, as detergent additives or for fabric desizing or
 CC starch liquefaction. They can also be used for the production of
 CC sweeteners and ethanol from starch. See also AAW14498-99.
 XX
 XX Sequence 478 AA;
 SQ
 Query Match 20.7%; Score 6; DB 17; Length 478;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 12 PTDSSEL 17
 Db 353 ptdsel 358
 |||||
 RESULT 31
 AAB19380
 ID AAB19380 standard; Protein; 489 AA.
 AC AAB19380;
 XX
 XX 06-MAR-2001 (first entry)
 DT
 DE Amino acid sequence of a human lipoxigenase protein.
 XX
 XX Human; lipoxigenase; leukotriene; lipid; chemotactic agent;
 KW inflammation; smooth muscle contraction; asthma; eye disease; arthritis;
 KW lung disease; cancer; acne; psoriasis.
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO2000061765-A2.
 PN
 XX
 XX 19-OCT-2000.
 PD
 XX
 XX 12-APR-2000; 2000WO-US09657.
 PF
 XX
 XX 12-APR-1999; 99US-0128817.
 PR
 XX 24-AUG-1999; 99US-0150454.
 XX
 XX (LEXI-) LEXICON GENETICS INC.
 PA
 XX
 XX Turner CA, Zambrowicz B, Nehls M, Friedrich G, Sands AT;
 PI
 XX
 XX WPI; 2000-665134/64.
 DR
 XX N-PSDB; AAC61748.
 DR
 XX Novel polynucleotides encoding human lipoxigenase proteins useful for
 PT producing transgenic animals preferably mouse .
 PT
 XX
 XX Claim 3; Page 60-61; 83pp; English.
 PS
 XX

CC AAB19379-92 represent novel human lipoxigenase proteins. Lipoxigenases
 CC oxidise lipids to produce leukotrienes. Leukotrienes bind cognate
 CC receptors and trigger biological effects. Leukotrienes influence a
 CC variety of biological processes, and can serve as, inter alia, potent
 CC chemotactic agents and mediators of inflammation, smooth muscle
 CC contractions, etc.. Lipoxigenases and leukotrienes are implicated in
 CC a variety of diseases and disorders, such as asthma, eye diseases,
 CC arthritis, lung disease, cancer, acne, psoriasis, etc..
 XX
 XX Sequence 489 AA;
 SQ
 Query Match 20.7%; Score 6; DB 21; Length 489;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 11 LPTDSE 16
 Db 380 lptdse 385
 |||||
 RESULT 32
 AAB84206
 ID AAB84206 standard; Protein; 498 AA.
 AC AAB84206;
 XX
 XX 06-AUG-2001 (first entry)
 DT
 DE Amino acid sequence of a fungamyl-like alpha-amylase.
 XX
 XX Fungamyl-like alpha-amylase; glucoamylase; dextrinisation; maltose;
 KW alcohol; starch; dough improver; brewing; starch liquefaction.
 KW
 XX Aspergillus oryzae.
 OS
 XX
 XX WO200134784-A1.
 PN
 XX
 XX 17-MAY-2001.
 PD
 XX
 XX 10-NOV-2000; 2000WO-DK00626.
 PF
 XX
 XX 10-NOV-1999; 99DK-0001617.
 PR
 XX
 XX (NOVO) NOVOZYMES AS.
 PA
 XX
 XX Bisgard-Frantzen H, Svendsen A, Pedersen S;
 PI
 XX
 XX WPI; 2001-367478/38.
 DR
 XX N-PSDB; AAF90208.
 DR
 XX
 XX New variant of Fungamyl-like alpha-amylase, useful for production of
 PT maltose syrups, includes mutations that improve stability against heat
 PT and acidic pH -
 PT
 XX
 XX Claim 1; Page 42-45; 49pp; English.
 PS
 XX
 XX The present sequence represents a fungamyl-like alpha-amylase. The
 CC specification describes variants of this fungamyl-like alpha-amylase,
 CC which have an alteration in one the amino acid regions 98-110, 150-160,
 CC 161-167, 280-288, 448-455, and 468-475. Each alteration is a deletion
 CC or substitution of an amino acid or an insertion of an amino acid
 CC downstream of a particular position. The variants retain alpha-amylase
 CC activity, and have better heat stability and/or stability at acidic pH,
 CC relative to wild-type enzyme. The variants can therefore be used at
 CC higher temperatures (more efficient conversion or faster reaction, and
 CC have reduced need for cooling and reduced risk of contamination). The
 CC variants may also be used in conjunction with other enzymes,
 CC particularly glucoamylase during dextrinisation. The variants are
 CC used to produce syrups, particularly of high maltose content, or alcohol,
 CC from starch; as dough improver for baked goods; in brewing, to increase
 CC fermentability of the wort; and for liquefaction of starch.
 CC
 XX

SQ Sequence 498 AA;

Query Match 20.7%; Score 6; DB 22; Length 498;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 PTSEL 17
 |||||
 Db 373 ptdsel 378

RESULT 33
 ABB67817
 ID ABB67817 standard; Protein; 503 AA.
 AC ABB67817;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 30243.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL11920.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 30243; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 503 AA;

Query Match 20.7%; Score 6; DB 22; Length 503;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 TDSELA 18
 |||||
 Db 308 tdsela 313

RESULT 34
 AAB19381
 ID AAB19381 standard; Protein; 556 AA.
 XX
 AC AAB19381;
 XX
 DT 06-MAR-2001 (first entry)
 XX
 DE Amino acid sequence of a human lipoxigenase protein.
 XX
 KW Human; lipoxigenase; leukotriene; lipid; chemotactic agent;
 KW inflammation; smooth muscle contraction; asthma; eye disease; arthritis;
 KW lung disease; cancer; acne; psoriasis.
 XX
 OS Homo sapiens.
 XX
 PN WO2000061765-A2.
 XX
 PD 19-OCT-2000.
 XX
 PF 12-APR-2000; 2000WO-US09657.
 XX
 PR 12-APR-1999; 99US-0128817.
 PR 24-AUG-1999; 99US-0150454.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 XX
 PI Turner CA, Zambrowicz B, Nehls M, Friedrich G, Sands AT;
 XX
 DR WPI; 2000-665134/64.
 DR N-PSDB; AAC61749.
 XX
 PT Novel polynucleotides encoding human lipoxigenase proteins useful for
 PT producing transgenic animals preferably mouse -
 XX
 PS Claim 3; Page 62-64; 83pp; English.
 XX
 CC AAB19379-92 represent novel human lipoxigenase proteins. Lipoxigenases
 CC oxidise lipids to produce leukotrienes. Leukotrienes bind cognate
 CC receptors and trigger biological effects. Leukotrienes influence a
 CC variety of biological processes, and can serve as, inter alia, potent
 CC chemotactic agents and mediators of inflammation, smooth muscle
 CC contractions, etc.. Lipoxigenases and leukotrienes are implicated in
 CC a variety of diseases and disorders, such as asthma, eye diseases,
 CC arthritis, lung disease, cancer, acne, psoriasis, etc..
 XX
 SQ Sequence 556 AA;

Query Match 20.7%; Score 6; DB 21; Length 556;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 LPTDSE 16
 |||||
 Db 225 lptdse 230

RESULT 35
 ABB71195
 ID ABB71195 standard; Protein; 602 AA.
 XX
 AC ABB71195;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 40377.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX

PN WO200171042-A2.
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US09231.
 PF
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL15298.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 40377; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell-signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 602 AA;

Query Match 20.7%; Score 6; DB 22; Length 602;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 LPTDSE 16
 Db 109 lptdse 114
 |||||

RESULT 36
 AAB19383
 ID AAB19383 standard; Protein; 615 AA.
 AC
 XX AAB19383;
 XX
 DT 06-MAR-2001 (first entry)
 XX
 DE Amino acid sequence of a human lipoxigenase protein.
 XX
 KW Human; lipoxigenase; leukotriene; lipid; chemotactic agent;
 KW inflammation; smooth muscle contraction; asthma; eye disease; arthritis;
 KW lung disease; cancer; acne; psoriasis.
 XX
 OS Homo sapiens.
 XX
 XX WO200061765-A2.
 PN
 XX
 PD 19-OCT-2000.
 XX
 PF 12-APR-2000; 2000WO-US09657.
 XX
 XX 12-APR-1999; 99US-0128817.
 PR
 PR 24-AUG-1999; 99US-0150454.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 XX

PI Turner CA, Zambrowicz B, Nehls M, Friedrich G, Sands AT;
 XX
 DR WPI; 2000-665134/64.
 DR N-PSDB; AAC61751.
 XX
 PT Novel polynucleotides encoding human lipoxigenase proteins useful for
 PT producing transgenic animals preferably mouse -
 XX
 PS Claim 3; Page 66-67; 83pp; English.
 XX
 CC AAB19379-92 represent novel human lipoxigenase proteins. Lipoxigenases
 CC oxidise lipids to produce leukotrienes. Leukotrienes bind cognate
 CC receptors and trigger biological effects. Leukotrienes influence a
 CC variety of biological processes, and can serve as, inter alia, potent
 CC chemotactic agents and mediators of inflammation, smooth muscle
 CC contractions, etc.. Lipoxigenases and leukotrienes are implicated in
 CC a variety of diseases and disorders, such as asthma, eye diseases,
 CC arthritis, lung disease, cancer, acne, psoriasis, etc..
 XX
 SQ Sequence 615 AA;

Query Match 20.7%; Score 6; DB 21; Length 615;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 LPTDSE 16
 Db 380 lptdse 385
 |||||

RESULT 37
 AAR86867
 ID AAR86867 standard; Protein; 616 AA.
 XX
 AC AAR86867;
 XX
 DT 27-AUG-1996 (first entry)
 XX
 DE Rat protocadherin pc5.
 XX
 KW Protocadherin; pc3; pc4; pc5; human; rat; cadherin; cell adhesion;
 KW catenin; therapy.
 XX
 OS Rattus rattus.
 XX
 PN WO9600289-A1.
 XX
 PD 04-JAN-1996.
 XX
 PF 26-JUN-1995; 95WO-US08071.
 XX
 PR 27-JUN-1994; 94US-0268161.
 XX
 PA (DOHE-) DOHENY EYE INST.
 XX
 PI Suzuki S;
 XX
 DR WPI; 1996-068873/07.
 DR N-PSDB; AAT03574.
 XX
 XX Polynucleotide(s) encoding human protocadherins pc3 and pc4 and rat
 PT pc5 - involved in cell-cell adhesion and regulation activities
 PT
 PS Claim 17; Page 130-132; 146pp; English.
 XX
 CC AAR86865-R86867 represent the sequences for three protocadherins. This
 CC sequence represents the rat protocadherin pc5. These sequences are
 CC related to cadherin, and possess cell adhesive ability. Cadherins are
 CC glycosylated integral membrane proteins that are involved in cell-cell
 CC adhesion. Cadherins are composed of an N-terminal extracellular domain
 CC which consists of 5 unique subdomains, a membrane spanning domain, and a
 CC C-terminal cytoplasmic domain. The cytoplasmic domain interacts with

CC the cytoskeleton through catenins and other cytoskeleton associated
CC proteins. The cytoplasmic domain is not present in all cadherins, but in
CC those which possess it, it is essential for the cadherins adhesive
CC function. The cadherins which do not possess a cytoplasmic domain appear
CC to function via a different method from those with a cytoplasmic domain.
CC These sequences were isolated using primers 1 and 2 (see AAR03575 and
CC AAR03576). The proteins may have regulatory functions in the cell, as
CC well as the cell-cell adhesive properties. Antibodies produced against
CC these sequences are useful for modulating the binding activity of these
CC protocadherins, and can be used therapeutically.

XX Sequence 616 AA;

Query Match 20.7%; Score 6; DB 17; Length 616;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 ELAPRS 21
Db 372 elaprs 377
|||||

RESULT 38
AAB19391
ID AAB19391 standard; Protein; 645 AA.

XX AC AAB19391;

XX DT 06-MAR-2001 (first entry)

XX DE Amino acid sequence of a human lipoxxygenase protein.

XX KW Human; lipoxxygenase; leukotriene; lipid; chemotactic agent;
XX inflammation; smooth muscle contraction; asthma; eye disease; arthritis;
XX lung disease; cancer; acne; psoriasis.

XX OS Homo sapiens.

XX PN WO200061765-A2.

XX PD 19-OCT-2000.

XX PF 12-APR-2000; 2000WO-US09657.

XX PR 12-APR-1999; 99US-0128817.

XX PR 24-AUG-1999; 99US-0150454.

XX PA (LEXI-) LEXICON GENETICS INC.

XX PI Turner CA, Zambrowicz B, Nehls M, Friedrich G, Sands AT;

XX DR WPI; 2000-665134/64.

XX DR N-PSDB; AAC61759.

XX PT Novel polynucleotides encoding human lipoxxygenase proteins useful for
XX producing transgenic animals preferably mouse

XX PS Claim 3; Page 77-79; 83pp; English.

XX CC AAB19379-92 represent novel human lipoxxygenase proteins. Lipoxxygenases
CC oxidise lipids to produce leukotrienes. Leukotrienes bind cognate
CC receptors and trigger biological effects. Leukotrienes influence a
CC variety of biological processes, and can serve as, inter alia, potent
CC chemotactic agents and mediators of inflammation, smooth muscle
CC contractions, etc.. Lipoxxygenases and leukotrienes are implicated in
CC a variety of diseases and disorders, such as asthma, eye diseases,
XX arthritis, lung disease, cancer, acne, psoriasis, etc..

XX Sequence 645 AA;

Query Match 20.7%; Score 6; DB 21; Length 645;

Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 LPTDSE 16
Db 536 lptdse 541
|||||

RESULT 39
ABB64989
ID ABB64989 standard; Protein; 704 AA.

XX AC ABB64989;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 21759.

XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL09092.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -

XX PS Disclosure; SEQ ID NO 21759; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).

XX CC The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 704 AA;

Query Match 20.7%; Score 6; DB 22; Length 704;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 PTDSEL 17
Db 176 ptdsel 181
|||||

RESULT 40
ABB66726
ID ABB66726 standard; Protein; 704 AA.
XX

AC ABB66726;
XX
XX DT 26-MAR-2002 (first entry)
XX
XX DE Drosophila melanogaster polypeptide SEQ ID NO 26970.
XX
XX KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
XX OS Drosophila melanogaster.
XX
XX PN WO200171042-A2.
XX
XX PD 27-SEP-2001.
XX
XX PF 23-MAR-2001; 2001WO-US09231.
XX
XX PR 23-MAR-2000; 2000US-191637P.
XX
XX PR 11-JUL-2000; 2000US-0614150.
XX
XX PA (PEKE) PE CORP NY.
XX
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX DR WPI; 2001-656860/75.
XX
XX DR N-PSDB; ABL10829.
XX
XX XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX PS Disclosure; SEQ ID NO 26970; 21pp + Sequence Listing; English.
XX
XX CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 704 AA;

Query Match 20.7%; Score 6; DB 22; Length 704;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSEL 17
DB 176 ptdsel 181
|||||

RESULT 41
AAB19379
ID AAB19379 standard; Protein; 711 AA.
XX
XX AC AAB19379;
XX
XX DT 06-MAR-2001 (first entry)
XX
XX DE Amino acid sequence of a human lipoxigenase protein.
XX
XX KW Human; lipoxigenase; leukotriene; lipid; chemotactic agent;
KW inflammation; smooth muscle contraction; asthma; eye disease; arthritis;
KW lung disease; cancer; acne; psoriasis.
XX
XX OS Homo sapiens.
XX

PN WO200061765-A2.
XX
XX PD 19-OCT-2000.
XX
XX PF 12-APR-2000; 2000WO-US09657.
XX
XX PR 12-APR-1999; 99US-0128817.
XX
XX PR 24-AUG-1999; 99US-0150454.
XX
XX PA (LEXI-) LEXICON GENETICS INC.
XX
XX PI Turner CA, Zambrowicz B, Nehls M, Friedrich G, Sands AT;
XX
XX DR WPI; 2000-665134/64.
XX
XX DR N-PSDB; AAC61747.
XX
XX PT Novel polynucleotides encoding human lipoxigenase proteins useful for
PT producing transgenic animals preferably mouse -
XX
XX PS Claim 3; Page 58-60; 83pp; English.
XX
XX CC AAB19379-92 represent novel human lipoxigenase proteins. Lipoxigenases
CC oxidise lipids to produce leukotrienes. Leukotrienes bind cognate
CC receptors and trigger biological effects. Leukotrienes influence a
CC variety of biological processes, and can serve as, inter alia, potent
CC chemotactic agents and mediators of inflammation, smooth muscle
CC contractions, etc.. Lipoxigenases and leukotrienes are implicated in
CC a variety of diseases and disorders, such as asthma, eye diseases,
CC arthritis, lung disease, cancer, acne, psoriasis, etc..
XX
XX SQ Sequence 711 AA;

Query Match 20.7%; Score 6; DB 21; Length 711;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 LPTDSE 16
DB 380 lptdse 385
|||||

RESULT 42
ABB69061
ID ABB69061 standard; Protein; 734 AA.
XX
XX AC ABB69061;
XX
XX DT 26-MAR-2002 (first entry)
XX
XX DE Drosophila melanogaster polypeptide SEQ ID NO 33975.
XX
XX KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
XX OS Drosophila melanogaster.
XX
XX PN WO200171042-A2.
XX
XX PD 27-SEP-2001.
XX
XX PF 23-MAR-2001; 2001WO-US09231.
XX
XX PR 23-MAR-2000; 2000US-191637P.
XX
XX PR 11-JUL-2000; 2000US-0614150.
XX
XX PA (PEKE) PE CORP NY.
XX
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX DR WPI; 2001-656860/75.
XX
XX DR N-PSDB; ABL13164.
XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX Disclosure; SEQ ID NO 33975; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 734 AA;

Query Match 20.7%; Score 6; DB 22; Length 734;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 LPTDSE 16
 |||||
 Db 363 lptdse 368

RESULT 43
 AAB19392
 ID AAB19392 standard; Protein; 771 AA.

XX AAB19392;

XX 06-MAR-2001 (first entry)

XX Amino acid sequence of a human lipoxigenase protein.

XX Human; lipoxigenase; leukotriene; lipid; chemotactic agent;
 KW inflammation; smooth muscle contraction; asthma; eye disease; arthritis;
 KW lung disease; cancer; acne; psoriasis.

XX Homo sapiens.

XX WO2000061765-A2.

XX 19-OCT-2000.

XX 12-APR-2000; 2000WO-US09657.

XX 12-APR-1999; 99US-0128817.

XX 24-AUG-1999; 99US-0150454.

XX (LEXI-) LEXICON GENETICS INC.

XX Turner CA, Zambrowicz B, Nehls M, Friedrich G, Sands AT;

XX WPI; 2000-665134/64.

XX N-PSDB; AAC61760.

XX Novel polynucleotides encoding human lipoxigenase proteins useful for
 PT producing transgenic animals preferably mouse -

XX Claim 3; Page 80-81; 83pp; English.

XX AAB19379-92 represent novel human lipoxigenase proteins. Lipoxigenases
 CC oxidise lipids to produce leukotrienes. Leukotrienes bind cognate
 CC receptors and trigger biological effects. Leukotrienes influence a
 CC variety of biological processes, and can serve as, inter alia, potent
 CC chemotactic agents and mediators of inflammation, smooth muscle
 CC contractions, etc.. Lipoxigenases and leukotrienes are implicated in

CC a variety of diseases and disorders, such as asthma, eye diseases,
 CC arthritis, lung disease, cancer, acne, psoriasis, etc..
 XX
 SQ Sequence 771 AA;

Query Match 20.7%; Score 6; DB 21; Length 771;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 LPTDSE 16
 |||||
 Db 536 lptdse 541

RESULT 44

AAU19687
 ID AAU19687 standard; Protein; 791 AA.

XX AAU19687;

XX 04-DEC-2001 (first entry)

XX Human novel extracellular matrix protein, Seq ID No 337.

XX Human; secreted extracellular matrix protein; immunomodulatory;
 KW Anti-HIV; antianemic; antirheumatic; antisclerotic; cardiant; vascular;
 KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;
 KW anti-atherosclerotic; immune/autoimmune disease; HIV infection; anaemia;
 KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;
 KW Scazary syndrome; Gaucher's disease; neurological diseases;
 KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
 KW cardiac arrest; tachycardia; angina; infection; corneal infections;
 KW wound healing; immunogen; gene therapy; antisense; food additive.

XX Homo sapiens.

XX WO200155368-A1.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01348.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 07-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 14-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220963.

XX 26-JUL-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225266.

XX 14-AUG-2000; 2000US-0225267.

XX 14-AUG-2000; 2000US-0225268.

XX 14-AUG-2000; 2000US-0225270.

XX 14-AUG-2000; 2000US-0225447.

XX 14-AUG-2000; 2000US-0225757.

XX 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 03-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0233223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 26-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 29-SEP-2000; 2000US-0236379.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.

PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465572/50.
DR N-PSDB; AAS31258.
XX
XX
PT Nucleic acid molecules encoding human secreted extracellular matrix
PT proteins, used in preventing, treating or ameliorating a disorder, e.g.
PT Alzheimer's and Parkinson's diseases and cancers -
XX
PS Claim 11; SEQ ID NO 337; 577pp; English.
XX

CC The invention relates to isolated nucleic acid molecules encoding
CC novel human secreted extracellular matrix proteins (SPs). The
CC polynucleotides and proteins are used to prevent, treat a medical
CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
CC chickens or sheep. For example, disorders associated with decreased
CC expression of SPs. The SP polynucleotide or a vector expressing them may
CC be administered to treat diseases by gene therapy. Antisense molecules
CC may be administered to down regulate expression of SPs by binding with
CC the cells own genes and preventing their expression. The polynucleotides
CC may also be used as DNA probes in diagnostic assays. The SPs may also be
CC used as antigens to produce antibodies and to identify modulators
CC (agonists and antagonists) of the SPs. The anti-(SP) antibodies and
CC antagonists may also be used to down regulate expression and activity of
CC SP and as diagnostic agents for detecting the presence of SPs in samples.
CC The disorders include for example: immune/autoimmune diseases (e.g. HIV
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
CC Parkinson's disease) cardio-/cerebrovascular disorders (e.g. cardiac
CC arrest, tachycardia and angina), infections caused by bacteria, viruses
CC and fungi and ocular disorders (e.g. corneal infections). Other uses
CC include wound healing, maintenance of organs before transplantation,

Query Match 20.7%; Score 6; DB 22; Length 791;

Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELAPRS 21
Db 435 elaprs 440

RESULT 45

ABE68766
ID ABE68766 standard; Protein; 816 AA.

XX AC ABE68766;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 33090.

XX KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL12869.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX PS Disclosure; SEQ ID NO 33090; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).

XX CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 816 AA;

Query Match 20.7%; Score 6; DB 22; Length 816;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDSELA 18

Db 28 tdsela 33

RESULT 46

ABB62835
ID ABB62835 standard; Protein; 830 AA.

XX

AC ABB62835;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 15297.

XX KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL06938.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX PS Disclosure; SEQ ID NO 15297; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).

XX CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 830 AA;

Query Match 20.7%; Score 6; DB 22; Length 830;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 DSELAP 19

Db 90 dselap 95

RESULT 47

ABG23820

ID ABG23820 standard; Protein; 832 AA.

XX AC ABG23820;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #23811.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.
XX XX
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI: 2001-639362/73.
XX DR N-PSDB; AAS88007.
XX XX
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -
XX PS Claim 20; SEQ ID NO 54179; 103pp; English.
XX XX
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG0010-ABG30377 represent novel human
XX CC diagnostic amino acid sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 832 AA;

Query Match 20.7%; Score 6; DB 22; Length 832;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 ELAPRS 21
| | | | |
Db 591 elaprs 596

RESULT 48
AAM79446
ID AAM79446 standard; Protein; 858 AA.
XX AC AAM79446;
XX DT 06-NOV-2001 (first entry)
XX XX Human protein SEQ ID NO 3092.
XX DE
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorder; arthritis; inflammation.
XX OS Homo sapiens.
XX XX

PN WO200157190-A2.
XX XX
XX PD 09-AUG-2001.
XX PF
XX PF 05-FEB-2001; 2001WO-US04098.
XX XX
XX PR 03-FEB-2000; 2000US-0496914.
XX PR 27-APR-2000; 2000US-0560875.
XX PR 20-JUN-2000; 2000US-0598075.
XX PR 19-JUL-2000; 2000US-0620325.
XX PR 01-SEP-2000; 2000US-0654936.
XX PR 15-SEP-2000; 2000US-0663561.
XX PR 20-OCT-2000; 2000US-0693325.
XX PR 30-NOV-2000; 2000US-0728422.
XX XX
XX PA (HYSE-) HYSEQ INC.
XX XX
XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
XX PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
XX PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX XX
XX DR WPI: 2001-476283/51.
XX DR N-PSDB; AAK52579.
XX XX
XX PT Nucleic acids encoding polypeptides with cytokine-like activities,
XX PT useful in diagnosis and gene therapy -
XX XX
XX PS Claim 20; Page 249-250; 6221pp; English.
XX XX
XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation.
XX CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
XX CC (AAM80020) are omitted as the relevant pages from the sequence listing
XX CC were missing at the time of publication.
XX SQ Sequence 858 AA;

Query Match 20.7%; Score 6; DB 22; Length 858;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 ELAPRS 21
| | | | |
Db 611 elaprs 616

RESULT 49
AAB19390
ID AAB19390 standard; Protein; 867 AA.
XX AC AAB19390;
XX XX
XX DT 06-MAR-2001 (first entry)
XX XX
XX DE Amino acid sequence of a human lipoxigenase protein.
XX KW Human; lipoxigenase; leukotriene; lipid; chemotactic agent;
XX KW inflammation; smooth muscle contraction; asthma; eye disease; arthritis;
XX KW lung disease; cancer; acne; psoriasis.
XX OS Homo sapiens.
XX XX
XX PN WO200061765-A2.

```
XX 19-OCT-2000.
PD 12-APR-2000; 2000WO-US09657.
PF 12-APR-1999; 99US-0128817.
XX 24-AUG-1999; 99US-0150454.
XX (LEXI-) LEXICON GENETICS INC.
PA Turner CA, Zambrowicz B, Nehls M, Friedrich G, Sands AT;
PI WPI; 2000-665134/64.
XX N-PSDB; AAC61758.
DR Novel polynucleotides encoding human lipoxigenase proteins useful for
PT producing transgenic animals preferably mouse -
XX Claim 3; Page 74-76; 83pp; English.
XX AAB19379-92 represent novel human lipoxigenase proteins. Lipoxigenases
CC oxidise lipids to produce leukotrienes. Leukotrienes bind cognate
CC receptors and trigger biological effects. Leukotrienes influence a
CC variety of biological processes, and can serve as, inter alia, potent
CC chemotactic agents and mediators of inflammation, smooth muscle
CC contractions, etc.. Lipoxigenases and leukotrienes are implicated in
CC a variety of diseases and disorders, such as asthma, eye diseases,
CC arthritis, lung disease, cancer, acne, psoriasis, etc..
XX SQ Sequence 867 AA;

Query Match 20.7%; Score 6; DB 21; Length 867;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 LPTDSE 16
Db 536 lptdse 541
|||||

RESULT 50
AA141743
ID AAY41743 standard; Protein; 916 AA.
AC AAY41743;
XX
XX 07-DEC-1999 (first entry)
DT
DE Human PRO707 protein sequence.
XX
KW Human; PRO: EST; expressed sequence tag; PCR primer; hybridisation;
KW probe; blood coagulation disorder; cancer; cellular adhesion disorder;
KW secreted protein; transmembrane protein.
XX
OS Homo sapiens.
XX
XX WO9946281-A2.
XX
PD 16-SEP-1999.
XX
XX 08-MAR-1999; 99WO-US05028.
XX
XX 10-MAR-1998; 98US-0077450.
PR 11-MAR-1998; 98US-0077632.
PR 11-MAR-1998; 98US-0077641.
PR 12-MAR-1998; 98US-0077649.
PR 13-MAR-1998; 98US-0077791.
PR 17-MAR-1998; 98US-0078004.
PR 20-MAR-1998; 98US-0040220.
PR 20-MAR-1998; 98US-0078886.
PR 20-MAR-1998; 98US-0078910.
PR 20-MAR-1998; 98US-0078936.
PR 20-MAR-1998; 98US-0078939.
PR 25-MAR-1998; 98US-0079294.
PR 26-MAR-1998; 98US-0079556.
PR 27-MAR-1998; 98US-0079663.
PR 27-MAR-1998; 98US-0079664.
PR 27-MAR-1998; 98US-0079689.
PR 27-MAR-1998; 98US-0079728.
PR 27-MAR-1998; 98US-0079786.
PR 30-MAR-1998; 98US-0079920.
PR 30-MAR-1998; 98US-0079923.
PR 31-MAR-1998; 98US-0080105.
PR 31-MAR-1998; 98US-0080107.
PR 31-MAR-1998; 98US-0080165.
PR 31-MAR-1998; 98US-0080194.
PR 01-APR-1998; 98US-0080327.
PR 01-APR-1998; 98US-0080328.
PR 01-APR-1998; 98US-0080333.
PR 01-APR-1998; 98US-0080334.
PR 08-APR-1998; 98US-0081049.
PR 08-APR-1998; 98US-0081070.
PR 08-APR-1998; 98US-0081071.
PR 09-APR-1998; 98US-0081195.
PR 09-APR-1998; 98US-0081203.
PR 09-APR-1998; 98US-0081229.
PR 15-APR-1998; 98US-0081817.
PR 15-APR-1998; 98US-0081838.
PR 15-APR-1998; 98US-0081952.
PR 15-APR-1998; 98US-0081955.
PR 21-APR-1998; 98US-0082568.
PR 21-APR-1998; 98US-0082569.
PR 22-APR-1998; 98US-0082700.
PR 22-APR-1998; 98US-0082704.
PR 22-APR-1998; 98US-0082804.
PR 23-APR-1998; 98US-0082767.
PR 23-APR-1998; 98US-0082796.
PR 27-APR-1998; 98US-0083336.
PR 28-APR-1998; 98US-0083322.
PR 29-APR-1998; 98US-0083392.
PR 29-APR-1998; 98US-0083495.
PR 29-APR-1998; 98US-0083496.
PR 29-APR-1998; 98US-0083499.
PR 29-APR-1998; 98US-0083500.
PR 29-APR-1998; 98US-0083545.
PR 29-APR-1998; 98US-0083554.
PR 29-APR-1998; 98US-0083558.
PR 30-APR-1998; 98US-0083559.
PR 30-APR-1998; 98US-0083742.
PR 05-MAY-1998; 98US-0084366.
PR 06-MAY-1998; 98US-0084414.
PR 06-MAY-1998; 98US-0084441.
PR 07-MAY-1998; 98US-0084598.
PR 07-MAY-1998; 98US-0084600.
PR 07-MAY-1998; 98US-0084640.
PR 07-MAY-1998; 98US-0084643.
PR 13-MAY-1998; 98US-0085323.
PR 13-MAY-1998; 98US-0085338.
PR 13-MAY-1998; 98US-0085339.
PR 15-MAY-1998; 98US-0085573.
PR 15-MAY-1998; 98US-0085579.
PR 15-MAY-1998; 98US-0085580.
PR 15-MAY-1998; 98US-0085582.
PR 15-MAY-1998; 98US-0085689.
PR 15-MAY-1998; 98US-0085697.
PR 15-MAY-1998; 98US-0085700.
PR 15-MAY-1998; 98US-0085704.
PR 18-MAY-1998; 98US-0086023.
PR 22-MAY-1998; 98US-0086392.
PR 22-MAY-1998; 98US-0086414.
PR 22-MAY-1998; 98US-0086430.
PR 22-MAY-1998; 98US-0086486.
```

PR 28-MAY-1998; 98US-0087098.
PR 28-MAY-1998; 98US-0087106.
PR 28-MAY-1998; 98US-0087208.
PR 30-JUL-1998; 98US-0094651.
PR 11-SEP-1998; 98US-0100038.
XX
PA (GETH) GENENTECH INC.
XX
PI Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
XX
DR WPI; 1999-551358/46.
DR N-PSDB; AAZ34221.
XX
PT New secreted and transmembrane polypeptides and their polynucleotides,
PT useful for treating blood coagulation disorders, cancers and cellular
PT adhesion disorders -
XX
PS Claim 12; Fig 157; 530pp; English.
XX
CC The present invention describes secreted and transmembrane polypeptides
CC and their polynucleotides. The nucleotide sequences are useful as
CC sources of probes, primers, for chromosome mapping, and for generation
CC of antisense sequences. They can also be used to create transgenic
CC animals. The proteins can be used to treat a variety of diseases and
CC disorders, depending on their function. Diseases that may be treated
CC include blood coagulation disorders, cancers and cellular adhesion
CC disorders. They may also be used to raise antibodies. AAZ33891 to
CC AAZ34338, and AAY41685 to AAY41774 represent polynucleotide and
CC polypeptide sequence given in the exemplification of the present
CC invention.
XX
SQ Sequence 916 AA;

Query Match 20.7%; Score 6; DB 20; Length 916;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 ELAPRS 21
Db 576 elaprs 581
|||||

Search completed: August 22, 2002, 11:19:15
Job time: 142 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: August 22, 2002, 11:16:53 ; Search time 13.5 Seconds
(without alignments)
83.175 Million cell updates/sec

Title: US-09-856-199-1

Perfect score: 29

Sequence: 1 KRCINQLLCKLPTDSEIAPRSXCHCRHL 29

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1200 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	72.4	381	1 SELP_HUMAN	P49908 homo sapien
2	9	31.0	380	1 SELP_MOUSE	P70274 mus musculus
3	9	31.0	385	1 SELP_RAT	P25236 rattus norv
4	9	31.0	402	1 SELP_BOVIN	P49907 bos taurus
5	7	24.1	559	1 PHO2_YEAST	P07269 saccharomyc
6	6	20.7	180	1 RSL1_ARATH	P10795 arabidopsis
7	6	20.7	205	1 KTHY_PYRHO	O59366 pyrococcus
8	6	20.7	302	1 YP93_RHILO	Q98kn6 rhizobium l
9	6	20.7	310	1 YAHB_ECOLI	P77700 escherichia
10	6	20.7	359	1 TPSB_CAEEL	Q20351 caenorhabdi
11	6	20.7	390	1 NCF1_MOUSE	Q09014 mus musculus
12	6	20.7	392	1 NCF1_BOVIN	O77774 bos taurus
13	6	20.7	394	1 MPK1_XENLA	Q05116 xenopus lae
14	6	20.7	416	1 LACY_KLEPN	P18817 klebsiella
15	6	20.7	498	1 AMYA_ASPAW	Q02905 aspergillus
16	6	20.7	499	1 AMYA_ASPOR	P10529 aspergillus
17	6	20.7	499	1 AMYB_ASPAW	Q02906 aspergillus
18	6	20.7	499	1 AMT_ASPSH	P30292 aspergillus
19	6	20.7	504	1 D8TP_SALTY	P40733 salmonella
20	6	20.7	622	1 SR68_CAEEL	Q20822 caenorhabdi
21	6	20.7	999	1 RLK5_ARATH	P47735 arabidopsis
22	6	20.7	1102	1 RPOP_AGART	P33539 agarius bi
23	6	20.7	2301	1 POLG_TMEVD	P13899 t genome po
24	6	20.7	2303	1 POLG_TMEVB	P08544 t genome po
25	6	20.7	2303	1 POLG_TMEVG	P28779 proteus vul
26	5	17.2	34	1 LPTN_PROVU	Q31282 bacterioph
27	5	17.2	53	1 VGK_BPAPH	P38038 bacterioph
28	5	17.2	53	1 VGK_BPAPH	O05755 mycobacteri
29	5	17.2	79	1 Y524_MYCLE	P03800 bacterioph
30	5	17.2	84	1 Y65_BPT7	P49304 borrelia bu
31	5	17.2	85	1 RS20_BORBU	Q10566 mycobacteri
32	5	17.2	87	1 Y898_MYCTU	P52866 daucus caro
33	5	17.2	106	1 RL36_DAUCA	

P02781	rattus norv	112	17.2	5	34	1	PSC2_RAT
O9m352	arabidopsis	112	17.2	5	35	1	R36B_ARATH
O80929	arabidopsis	113	17.2	5	36	1	R36A_ARATH
O14905	homo sapien	120	17.2	5	37	1	WN15_HUMAN
P47301	mycoplasma	123	17.2	5	38	1	Y055_MYCGE
P46984	saccharomyc	123	17.2	5	39	1	YJ54_YEAST
P16108	polyandroca	125	17.2	5	40	1	LECC_POLMI
P46744	prototheca	125	17.2	5	41	1	RT12_PROWI
Q923e6	xanthomonas	127	17.2	5	42	1	RL17_XANCP
P13303	bacterioph	128	17.2	5	43	1	Y050_BPT4
Q65694	bovine resp	136	17.2	5	44	1	VNS1_BRSPA
P44222	haemophilus	139	17.2	5	45	1	YE98_HAEIN
O67528	aquifex aeo	142	17.2	5	46	1	NDK_AQUAE
P37903	escherichia	144	17.2	5	47	1	UP03_ECOLI
Q03579	saccharomyc	150	17.2	5	48	1	Y08W_YEAST
Q49117	methyllobact	153	17.2	5	49	1	YMEC_METEX
Q9xip2	pseudomonas	159	17.2	5	50	1	NUDH_PSEAE
P26564	medicago sa	161	17.2	5	51	1	IF5A_MEDSA
P49085	bacterioph	164	17.2	5	52	1	ATPX_CYPAP
P39491	cyathophora	164	17.2	5	53	1	Y07E_BPT4
P41053	escherichia	165	17.2	5	54	1	YGAD_ECOLI
P28014	medicago sa	167	17.2	5	55	1	TCTP_MEDSA
P76483	escherichia	167	17.2	5	56	1	YFBW_ECOLI
P57701	pseudomonas	169	17.2	5	57	1	DSB2_PSEAE
O85290	buchnera ap	171	17.2	5	58	1	DAPD_BUCAP
Q9kv17	vibrio chol	181	17.2	5	59	1	ORN_VIBCH
Q37875	bacterioph	185	17.2	5	60	1	LYCV_BPPI
Q17819	caenorhabdi	193	17.2	5	61	1	ORN_CAEEL
Q15643	homo sapien	197	17.2	5	62	1	TRIA_HUMAN
P21466	bacillus su	199	17.2	5	63	1	RS4_BACSU
P57666	streptomyce	200	17.2	5	64	1	ORN_STRCO
Q9p1j1	campylobact	201	17.2	5	65	1	BIOD_CAMJE
P57667	streptomyce	201	17.2	5	66	1	ORN_STRGR
P07596	hordeum vul	203	17.2	5	67	1	IAAS_HORVU
Q08746	saccharomyc	203	17.2	5	68	1	Y08O_YEAST
P30817	simian herp	208	17.2	5	69	1	US10_HSVSB
P50870	enterococcu	209	17.2	5	70	1	SATA_ENTFC
Q9k6g5	bacillus ha	209	17.2	5	71	1	UPP_BACHD
Q9p3p0	ureaplasma	213	17.2	5	72	1	KAD_UREPA
P49202	bos taurus	215	17.2	5	73	1	UR2R_BOVIN
O87120	actinobacil	222	17.2	5	74	1	CDTA_ACTAC
O65522	haemophilus	223	17.2	5	75	1	CDTA_HAEDU
P42464	lycopersico	224	17.2	5	76	1	AGL9_LYCCE
P47649	mycoplasma	225	17.2	5	77	1	Y409_MYCGE
P33999	escherichia	225	17.2	5	78	1	YJUG_ECOLI
P23139	rhodospiril	238	17.2	5	79	1	YPE2_RHORI
P42378	lymaea sta	239	17.2	5	80	1	FRIY_LYMST
O32488	enterobacte	241	17.2	5	81	1	PHOU_ENTCL
Q03474	vibrio para	242	17.2	5	82	1	LAFS_VIBPA
P42451	synecchococ	243	17.2	5	83	1	SUMT_SYNP7
Q9zdn8	rickettsia	243	17.2	5	84	1	Y289_RICPR
P33312	saccharomyc	244	17.2	5	85	1	R187_YEAST
P10399	gallus gall	246	17.2	5	86	1	POLI_CHICK
O67610	aquifex aeo	248	17.2	5	87	1	FARG_AQUAE
Q06845	halobacteri	248	17.2	5	88	1	RL4_HALHA
Q00997	herpesvirus	249	17.2	5	89	1	VIEG_HSVSA
Q14017	schizosacch	249	17.2	5	90	1	YDPE_SCHPO
Q99550	mus musculus	257	17.2	5	91	1	GHSR_MOUSE
P87568	canine aden	259	17.2	5	92	1	E434_ADECT
P42643	arabidopsis	267	17.2	5	93	1	LA31_ARATH
P29972	homo sapien	269	17.2	5	94	1	AQPI_HUMAN
Q02013	mus musculus	269	17.2	5	95	1	AQPI_MOUSE
P29975	rattus norv	269	17.2	5	96	1	AQPI_RAT
P47865	bos taurus	271	17.2	5	97	1	AQPI_BOVIN
P56401	ovis aries	272	17.2	5	98	1	AQPI_SHEEP
P57323	buchnera ap	274	17.2	5	99	1	DAPD_BUCAI
Q97np1	streptococ	275	17.2	5	100	1	PURR_STRPN
Q59767	rhodospiril	277	17.2	5	101	1	NADB_RHORI
P21061	vaccinia vi	277	17.2	5	102	1	VA38_VACCC
P24763	vaccinia vi	277	17.2	5	103	1	VA38_VACCV
P37853	variola vir	277	17.2	5	104	1	VA38_VARV
O26151	methanobact	277	17.2	5	105	1	Y045_METTH
P19932	escherichia	285	17.2	5	106	1	HYAF_ECOLI

107	5	17.2	290	1	YOAU_BACSU	034701	180	5	17.2	416	1	CCA_HAEIN	P45269	haemophilus
108	5	17.2	292	1	HEMK_HAEIN	P45253	181	5	17.2	416	1	ILIS_RAT	P43303	rattus norv
109	5	17.2	294	1	Y237_MYCPN	P75455	182	5	17.2	420	1	IE68_HSV11	P04485	herpes simp
110	5	17.2	296	1	YA23_VIBCH	Q9kt82	183	5	17.2	423	1	IVD_HUMAN	P26440	homo sapien
111	5	17.2	297	1	RRP7_YEAST	P25368	184	5	17.2	424	1	SAMB_SALTY	P23832	salmonella
112	5	17.2	297	1	SMF_BACSU	P39813	185	5	17.2	425	1	TANK_HUMAN	Q92844	homo sapien
113	5	17.2	297	1	YMY9_YEAST	Q03161	186	5	17.2	425	1	KOTA_CHLPN	Q42222	chlamydia p
114	5	17.2	299	1	PFYF_CRIGR	Q60415	187	5	17.2	437	1	NQRA_VIBHA	Q91fw1	vibrio harv
115	5	17.2	299	1	YXFE_HUMAN	P40855	188	5	17.2	446	1	ACHA_BRARE	Q98880	brachydanio
116	5	17.2	299	1	YXFE_HUMAN	P75826	189	5	17.2	456	1	ACHA_CHICK	Q09479	gallus gall
117	5	17.2	300	1	NOCR_AGR75	Q00678	190	5	17.2	456	1	PURL_ARCFU	Q25388	archaeoglob
118	5	17.2	300	1	NOCR_AGR75	P35112	191	5	17.2	456	1	TRME_SYNY3	P73839	synecocyst
119	5	17.2	302	1	TEPI_RAT	Q02445	192	5	17.2	457	1	ACH1_XENLA	P24456	xenopus lae
120	5	17.2	303	1	VG49_HSVSA	Q01032	193	5	17.2	457	1	ACH2_XENLA	P03377	xenopus lae
121	5	17.2	305	1	Y452_BUCAT	P57527	194	5	17.2	457	1	ACHA_BOVIN	P02709	bos taurus
122	5	17.2	306	1	TEPI_MOUSE	O54819	195	5	17.2	457	1	ACHA_MOUSE	P04756	mus musculus
123	5	17.2	307	1	TRUB_PASMU	Q9cmq7	196	5	17.2	457	1	ACHA_RAT	P25108	rattus norv
124	5	17.2	311	1	GYP7_YARLI	P09379	197	5	17.2	461	1	ACHA_TORCA	P02710	torpedo cal
125	5	17.2	313	1	YA71_METJA	Q58471	198	5	17.2	461	1	ACHA_TORCA	P02710	torpedo cal
126	5	17.2	316	1	KHSE_MYCTU	Q10603	199	5	17.2	461	1	SELA_HAEIN	P43910	haemophilus
127	5	17.2	318	1	ANX4_CANFA	P50994	200	5	17.2	462	1	CISY_TETH	P24118	tetrahymena
128	5	17.2	321	1	ISP_BACCS	P29140	201	5	17.2	463	1	CMF1_BACSU	P39145	bacillus su
129	5	17.2	324	1	NU1M_CROLA	P34186	202	5	17.2	463	1	YAD4_YEAST	P28003	saccharomyc
130	5	17.2	327	1	FIP1_YEAST	P45976	203	5	17.2	464	1	BIOA_METJA	Q58696	methanococc
131	5	17.2	329	1	FAC2_HUMAN	Q9y256	204	5	17.2	465	1	YHJA_ECOLI	P37197	escherichia
132	5	17.2	329	1	FAC2_MOUSE	P57791	205	5	17.2	467	1	FUMC_ECOLI	P05042	escherichia
133	5	17.2	330	1	LDH_SCHPO	Q9p7p7	206	5	17.2	467	1	PHYA_ASPAW	P34753	aspergillus
134	5	17.2	333	1	GRP8_HUMAN	P48146	207	5	17.2	467	1	PHYA_ASPNG	P34752	aspergillus
135	5	17.2	337	1	ATH2_MOUSE	P48986	208	5	17.2	469	1	TRME_SYNEL	Q9x9t0	synecococc
136	5	17.2	339	1	MBR1_YEAST	P23493	209	5	17.2	470	1	YCJX_HAEIN	P44280	haemophilus
137	5	17.2	343	1	YD57_METJA	Q58752	210	5	17.2	475	1	GUNA_CLOCE	P17901	clostridium
138	5	17.2	351	1	LPXD_BRUME	Q44630	211	5	17.2	478	1	HEP2_MOUSE	P49182	mus musculus
139	5	17.2	355	1	PURK_ECOLI	P09029	212	5	17.2	479	1	HEP2_RAT	Q64268	rattus norv
140	5	17.2	358	1	LIP_PSEGL	Q05489	213	5	17.2	480	1	HEP2_RAT	P47776	oryctolagus
141	5	17.2	358	1	Y993_METJA	Q58400	214	5	17.2	481	1	KPYK_BUCAT	P57404	buchnera ap
142	5	17.2	359	1	GPDA_DROKA	Q97463	215	5	17.2	481	1	SESL_XENLA	P58003	xenopus lae
143	5	17.2	359	1	LACH_DROME	Q24372	216	5	17.2	482	1	ACHA_HUMAN	P02708	homo sapien
144	5	17.2	359	1	VAL1_TYLCM	P27260	217	5	17.2	492	1	SESL_HUMAN	Q9y6p5	homo sapien
145	5	17.2	359	1	VAL1_TYLCU	P38609	218	5	17.2	492	1	SESL_MOUSE	P58006	mus musculus
146	5	17.2	361	1	RPL_BUCAT	P57268	219	5	17.2	493	1	ACH2_CABEL	P48181	caenorhabdi
147	5	17.2	362	1	PURK_HAEIN	P43850	220	5	17.2	493	1	PUR1_SYNP7	Q55038	synecococc
148	5	17.2	364	1	GHSR_RAT	O08725	221	5	17.2	497	1	DEGP_CHLMU	Q9p197	chlamydia m
149	5	17.2	364	1	IHA_PIG	P04087	222	5	17.2	497	1	DEGP_CHLTR	P18584	chlamydia t
150	5	17.2	364	1	LIP_BURCE	P22088	223	5	17.2	499	1	RUBB_SECE	Q43831	secale cere
151	5	17.2	364	1	LIP_PSES5	P25275	224	5	17.2	500	1	ICL_HUMAN	P05155	homo sapien
152	5	17.2	366	1	GHSR_HUMAN	Q92847	225	5	17.2	504	1	ICL_MOUSE	P37290	mus musculus
153	5	17.2	366	1	GHSR_PIG	Q95254	226	5	17.2	509	1	CHLB_MESVI	Q9mur9	mesostigma
154	5	17.2	369	1	LMXB_MESAU	Q60564	227	5	17.2	512	1	V233_FOWPV	Q91502	owlpox vir
155	5	17.2	370	1	VP40_BDV	Q01552	228	5	17.2	512	1	HMSH_DROME	Q03372	drosophila
156	5	17.2	371	1	YNEK_ECOLI	P76150	229	5	17.2	515	1	GPBD_CANAL	O13287	candida alb
157	5	17.2	372	1	LMXB_MOUSE	O88609	230	5	17.2	517	1	SYM_MYCTU	O05593	mycobacteri
158	5	17.2	375	1	RND_ECOLI	P09155	231	5	17.2	519	1	CNA2_YEAST	P06776	saccharomyc
159	5	17.2	376	1	YQCB_ERWCA	Q47417	232	5	17.2	526	1	GCR2_YEAST	Q01722	saccharomyc
160	5	17.2	377	1	LMXB_CHICK	P53413	233	5	17.2	534	1	PCPB_FLAS3	P42535	flavobacter
161	5	17.2	377	1	YBHS_ECOLI	P75775	234	5	17.2	537	1	QAY_NEUCR	P11636	neurospora
162	5	17.2	378	1	LUXE_VIBHA	P14286	235	5	17.2	538	1	NADB_PSEAE	Q51363	pseudomonas
163	5	17.2	378	1	PDXB_ECOLI	P05459	236	5	17.2	540	1	ECM1_HUMAN	Q16610	homo sapien
164	5	17.2	379	1	LMXB_HUMAN	O60663	237	5	17.2	540	1	NADB_ECOLI	P10902	escherichia
165	5	17.2	379	1	PURK_BACSU	P12045	238	5	17.2	545	1	CP10_LYMTS	P48416	lymnaea sta
166	5	17.2	379	1	WZA_ECOLI	P76388	239	5	17.2	545	1	HGT1_CANAL	O74713	candida alb
167	5	17.2	381	1	NIEV_KLEPN	P05345	240	5	17.2	553	1	NADB_SYNY3	P74562	synecocyst
168	5	17.2	382	1	DUT_MCMVS	Q09151	241	5	17.2	571	1	PEN3_ADE02	P03276	human adeno
169	5	17.2	382	1	LMXA_MESAU	Q04650	242	5	17.2	571	1	TERM_ADE05	P12538	human adeno
170	5	17.2	383	1	Y944_SYNY3	P74311	243	5	17.2	575	1	GRK6_HUMAN	Q64752	avian adeno
171	5	17.2	388	1	ADH1_CLOSA	P13604	244	5	17.2	576	1	YD49_MYCTU	P43250	homo sapien
172	5	17.2	388	1	PURK_SYNY3	P74724	245	5	17.2	579	1	GPC3_HUMAN	Q11019	mycobacteri
173	5	17.2	388	1	SOPA_ECOLI	P08866	246	5	17.2	580	1	RECN_SYNY3	P51654	homo sapien
174	5	17.2	390	1	NCFL_HUMAN	P14598	247	5	17.2	584	1	CYA3_RHIME	P74374	synecocyst
175	5	17.2	395	1	ACKA_BACSU	P37877	248	5	17.2	587	1	RUBB_BRANA	Q9z330	rhizobium m
176	5	17.2	395	1	NEUA_STRAG	Q53598	249	5	17.2	588	1	GV7_XENLA	P21241	brassica na
177	5	17.2	397	1	TRPB_BUCDN	O68428	250	5	17.2	591	1	RUBB_PEA	P20398	xenopus lae
178	5	17.2	405	1	ACKA_BORBU	O51567	251	5	17.2	595	1	GPC3_RAT	P08927	pisum sativ
179	5	17.2	414	1	YM3_STRCO	P14707	252	5	17.2	597	1		P13265	rattus norv

253 1 RUBB_ARATH 600
 254 1 FET1_CHICK 606
 255 1 FET1_HORSE 609
 256 1 F1MB_DICDI 610
 257 1 FET1_RAT 611
 258 1 HTPG_HELPJ 621
 259 1 HTPG_HELPJ 621
 260 1 Y017_RICPR 627
 261 1 KPC3_DROME 634
 262 1 SC1_RAT 634
 263 1 YA2A_SCHPO 642
 264 1 HM38_CABEL 644
 265 1 RN17_HUMAN 653
 266 1 RNB_HAEN 659
 267 1 VST2_HEVME 659
 268 1 PHUB_ECOLI 660
 269 1 FTSB_CAPAN 662
 270 1 YME1_SCHMA 662
 271 1 MNE1_YEAST 663
 272 1 Y366_MYCGE 667
 273 1 YKCA_CABEL 668
 274 1 CACP_YEAST 670
 275 1 PPCK_CUCSA 670
 276 1 LOX5_RAT 672
 277 1 LOX5_MOUSE 673
 278 1 MTF1_MOUSE 675
 279 1 OR1_COTJA 676
 280 1 YMEH_CABEL 676
 281 1 CA02_RAT 681
 282 1 RPOC_ORISA 682
 283 1 RPOC_MAIZE 683
 284 1 RPOC_WHEAT 683
 285 1 PHUB_SALTY 685
 286 1 HS80_LYCES 685
 287 1 HS82_ARATH 689
 288 1 HS82_ORISA 699
 289 1 HS83_ARATH 699
 290 1 HS81_ARATH 700
 291 1 AMYG_CLOSP 702
 292 1 HS83_PHANI 703
 293 1 FTH2_ARATH 704
 294 1 FTSB_MEDSA 706
 295 1 ORC1_SCHPO 707
 296 1 FTSB_TOBAC 714
 297 1 GP85_TRYCR 714
 298 1 HS82_MAIZE 715
 299 1 PSAA_HOPSQ 715
 300 1 FTH1_ARATH 716
 301 1 PSAA_ARAAA 716
 302 1 PSAA_CYCRE 717
 303 1 EXT2_MOUSE 718
 304 1 GROU_DROME 719
 305 1 MCM7_HUMAN 719
 306 1 MCM7_MOUSE 719
 307 1 PSAA_ADICA 719
 308 1 PSAA_ASPND 719
 309 1 PSAA_ENCLE 719
 310 1 PSAA_EOUPA 719
 311 1 PSAA_PSIHU 719
 312 1 PSAA_TORCL 719
 313 1 MCM7_XENIA 720
 314 1 PSAA_EPHTW 720
 315 1 PSAA_MARBO 720
 316 1 PSAA_WELMI 720
 317 1 PSAA_ANGVE 721
 318 1 PSAA_GINBI 721
 319 1 PRTP_PRIVF 724
 320 1 MK10_YEAST 733
 321 1 YH09_YEAST 734
 322 1 EZH1_HUMAN 747
 323 1 EZH1_MOUSE 747
 324 1 PSAA_MARPO 750
 325 1 PSAA_PINTH 753

P21240 arabidopsis
 Q90922 gallus gall
 P49066 equus caball
 P54680 dictyosteli
 P02773 rattus norv
 Q92mm2 helicobacte
 P56116 helicobacte
 Q92ec6 rickettsia
 P13678 drosophila
 P24084 rattus norv
 Q09701 schizosacch
 Q19720 caenorhabdi
 Q9bxt8 homo sapien
 P44440 haemophilus
 Q03500 hepatitis e
 P06972 escherichia
 Q39444 capsicum an
 P46508 schistosoma
 P24720 saccharomyc
 P47606 mycoplasma
 P42083 caenorhabdi
 P32796 saccharomyc
 P42066 cucumis sat
 P12527 rattus norv
 P48999 mus musculu
 Q07243 mus musculu
 P23499 coturnix co
 P54813 caenorhabdi
 P97562 rattus norv
 P12092 oryza sativ
 P16024 zea mays (m
 Q9xps8 triticum ae
 O87656 salmonella
 P36181 lycopersico
 P55737 arabidopsis
 P33126 oryza sativ
 P51818 arabidopsis
 P27323 arabidopsis
 P29761 clostridium
 P51819 pharbitis n
 Q9fh02 arabidopsis
 Q9bae0 medicago sa
 P54789 schizosacch
 O82150 nicotiana t
 Q03877 trypanosoma
 Q08277 zea mays (m
 Q9mu11 huperzia sq
 Q39102 arabidopsis
 Q9mu17 araucaria a
 Q9mu10 cycas revol
 P70428 mus musculu
 P16371 drosophila
 P33993 homo sapien
 Q61881 mus musculu
 Q9mu13 adiantum ca
 Q9mu14 asplenium n
 Q9mu16 encephalart
 Q9mu16 equisetum p
 Q9mu12 psilotum nu
 Q9mu10 torreya cal
 Q91876 xenopus lae
 Q9mu18 ephedra twe
 Q9mu11 marsilea bo
 Q9mu12 welwitschia
 Q9mu15 angiopteris
 Q9mu10 ginkgo bilo
 P11871 pseudorabie
 Q02197 saccharomyc
 P32900 saccharomyc
 Q92800 homo sapien
 P70351 mus musculu
 P06406 marchantia
 P41639 pinus thunb

P28944 equine herp
 P32479 saccharomyc
 P56726 mus musculu
 P97698 rattus norv
 Q50288 mycoplasma
 P37868 salmonella
 P21564 torpedo mar
 O07451 alcaligenes
 Q9wv8 mus musculu
 P37717 acetobacter
 P35522 torpedo cal
 P47046 saccharomyc
 P18824 drosophila
 Q9nsv4 homo sapien
 P09186 glycine max
 P34595 caenorhabdi
 Q9jlf7 mus musculu
 P24095 glycine max
 P38414 lens culina
 P29044 barley yell
 P29045 barley yell
 O78441 guillardia
 Q61361 mus musculu
 P50668 rattus norv
 Q50654 mycobacteri
 Q97n28 clostridium
 Q28062 bos taurus
 P25389 saccharomyc
 P32790 homo sapien
 P54394 bacillus su
 O14939 homo sapien
 P97813 mus musculu
 P70498 rattus norv
 O02668 sus scrofa
 Q9wrl5 herpesvirus
 Q03631 saccharomyc
 P19823 homo sapien
 P97279 mesocricetu
 Q61703 mus musculu
 Q24119 drosophila
 Q47952 haemophilus
 Q47952 haemophilus
 P73505 synechocyst
 P06695 pseudomonas
 P51565 escherichia
 Q43303 homo sapien
 Q60432 cricetus
 P23378 homo sapien
 P6156 vibrio furn
 P47047 saccharomyc
 P35384 bos taurus
 P17468 bovine rota
 P21615 bovine rota
 P22678 simian 11 r
 P15436 saccharomyc
 Q9smv7 arabidopsis
 P09361 mus musculu
 O70361 mus musculu
 P13595 mus musculu
 Q38953 arabidopsis
 P40541 lymphaea sta
 Q61309 lymphaea sta
 Q9b76 homo sapien
 Q9n1v1 drosophila
 Q61087 mus musculu
 Q92207 mus musculu
 Q13751 homo sapien
 P58402 escherichia
 P30855 escherichia
 Q10947 caenorhabdi
 P08546 human cytomc
 P15801 saccharomyc
 Q24020 drosophila
 P33345 escherichia

399	5	17.2	1337	1	YDM5_SCHPO	P87136 schizosacch	472	4	13.8	71	1	YE34_ANASP	Q4148 anabaena sp
400	5	17.2	1425	1	SWS_DROME	Q9u969 drosophila	473	4	13.8	72	1	NRDH_LACLC	Q48708 lactococcus
401	5	17.2	1426	1	CUT2_MOUSE	P70298 mus musculu	474	4	13.8	73	1	CSRA_TREPA	Q83663 treponema p
402	5	17.2	1482	1	NME2_MOUSE	Q01097 mus musculu	475	4	13.8	73	1	YG89_ARCFU	Q28584 archaeoglob
403	5	17.2	1482	1	NME2_RAT	Q00960 rattus norv	476	4	13.8	74	1	EDDF_HUMAN	Q60584 homo sapien
404	5	17.2	1484	1	NME2_HUMAN	Q13224 rattus norv	477	4	13.8	75	1	EX7S_BACSU	P54322 bacillus su
405	5	17.2	1524	1	Y133_HUMAN	Q14146 homo sapien	478	4	13.8	75	1	IPKA_CHICK	Q90841 gallus gall
406	5	17.2	1538	1	PPSB_MYCTU	Q10978 mycobacteri	479	4	13.8	75	1	RPOZ_MESVI	Q9mul7 mesostigma
407	5	17.2	1548	1	UGG3_DROME	Q09332 drosophila	480	4	13.8	75	1	VLYS_BPM2	P03609 bacterioph
408	5	17.2	1595	1	SOS_DROME	P26675 drosophila	481	4	13.8	76	1	CX02_CONTE	Q9x2k9 conus texti
409	5	17.2	1668	1	DPOL_THEHY	Q9hn05 thermococcu	482	4	13.8	77	1	CXK1_CONTE	P18512 conus texti
410	5	17.2	1679	1	YMF9_YEAST	Q04958 saccharomyc	483	4	13.8	78	1	EX7S_MYCLE	Q9x784 mycobacteri
411	5	17.2	1713	1	LMA3_HUMAN	Q16787 homo sapien	484	4	13.8	78	1	YO09_BPL2	P42544 bacterioph
412	5	17.2	1713	1	PMPC_CHLTR	O84419 chlamydia t	485	4	13.8	80	1	CALM_STRPU	P05934 strongyloce
413	5	17.2	1770	1	AKAC_HUMAN	Q02952 homo sapien	486	4	13.8	80	1	Y14C_BPT4	P39506 bacterioph
414	5	17.2	1781	1	ITB4_RAT	Q64632 rattus norv	487	4	13.8	80	1	YC15_OENV1	P15566 oenothera v
415	5	17.2	1807	1	MYH2_HUMAN	Q9ukx2 homo sapien	488	4	13.8	81	1	FER_BACSU	P50727 bacillus su
416	5	17.2	1941	1	CIN5_HUMAN	Q14524 homo sapien	489	4	13.8	81	1	FER_BACTH	P10245 bacillus th
417	5	17.2	2019	1	CIN5_RAT	P15389 rattus norv	490	4	13.8	81	1	YQ72_BACAN	Q9rmw2 bacillus an
418	5	17.2	2146	1	INSR_DROME	P09208 drosophila	491	4	13.8	83	1	GVG1_HALN1	P24371 halobacteri
419	5	17.2	2248	1	CYA1_DROME	P32870 drosophila	492	4	13.8	83	1	GVG2_HALME	P33960 halobacteri
420	5	17.2	2424	1	CCAA_RABIT	P27884 oryctolagus	493	4	13.8	83	1	GVPG_HALME	Q02232 halobacteri
421	5	17.2	2493	1	CYAA_USTMA	P49606 ustilago ma	494	4	13.8	84	1	RS14_ARATH	P24036 arabidopsis
422	5	17.2	2688	1	ZEP1_MOUSE	Q03172 mus musculu	495	4	13.8	84	1	YFAE_ECOLI	P37910 escherichia
423	5	17.2	2787	1	TEL1_YEAST	P38110 saccharomyc	496	4	13.8	84	1	YOAF_ECOLI	P76244 escherichia
424	5	17.2	2895	1	HYDP_DROME	P51592 drosophila	497	4	13.8	85	1	CPB1_ECOLI	P13959 escherichia
425	5	17.2	3164	1	POLG_TUMVJ	P89509 t genome po	498	4	13.8	85	1	YMI5_PARTE	P15616 paramecium
426	5	17.2	3255	1	POLG_LMV0	P31999 l genome po	499	4	13.8	86	1	CPB3_ECOLI	P03855 escherichia
427	5	17.2	3255	1	POLG_LMVE	P89876 l genome po	500	4	13.8	86	1	PTHP_BORBU	O51507 borrelia bu
428	5	17.2	3655	1	YAMB_SCHPO	Q10064 schizosacch	501	4	13.8	86	1	SODC_PASHA	Q51853 pasteurella
429	5	17.2	3830	1	SACS_MOUSE	Q9j1c8 mus musculu	502	4	13.8	87	1	A4_MACFA	P53601 macaca fasc
430	5	17.2	3898	1	POLG_BVDVS	Q01499 bovine vira	503	4	13.8	87	1	HCYA_TACTR	P02243 tachypleus
431	5	17.2	4725	1	DYHC_DICDI	P34036 dictyosteli	504	4	13.8	87	1	RR15_TOBAC	P05973 nicotiana t
432	5	17.2	5147	1	FAT_DROME	P33450 drosophila	505	4	13.8	87	1	SODC_ACTAC	O59081 actinobacil
433	4	13.8	17	1	RM35_YEAST	P36530 saccharomyc	506	4	13.8	88	1	HIPB_ECOLI	P23873 escherichia
434	4	13.8	24	1	DMS6_PHYBI	P81490 phyllomedus	507	4	13.8	88	1	RR15_ARATH	P56805 arabidopsis
435	4	13.8	26	1	SMS1_AMICA	Q9prz6 amia calva	508	4	13.8	88	1	SSAS_SALTY	P74891 salmonella
436	4	13.8	27	1	CD59_SHEEP	P58020 ovine aries	509	4	13.8	88	1	VSPM_CVMJH	P06591 murine coro
437	4	13.8	30	1	ANF_RANRI	P09196 rana ridibu	510	4	13.8	90	1	HS72_USTMA	P18694 ustilago ma
438	4	13.8	30	1	CALM_LYTP1	P05935 lytechinus	511	4	13.8	90	1	YANF_SCHPO	Q10189 schizosacch
439	4	13.8	34	1	DEF2_RABIT	P07468 oryctolagus	512	4	13.8	91	1	CP83_BF186	P41061 bacterioph
440	4	13.8	35	1	SBST_PINTH	P41625 pinus thunb	513	4	13.8	92	1	B2MG_MUSCR	P55077 mus caroli
441	4	13.8	37	1	RK36_ASTLO	P24355 astasia lon	514	4	13.8	92	1	IAPP_CAVPO	P21266 cavia porce
442	4	13.8	38	1	SCR7_PANIM	P58490 pandinus lm	515	4	13.8	92	1	MCI_PINRA	Q24493 pinus radia
443	4	13.8	40	1	BD10_BOVIN	P46168 bos taurus	516	4	13.8	92	1	Y10B_BPT4	P22278 bacterioph
444	4	13.8	41	1	PSBK_CYACA	Q9t1r0 cyanidium c	517	4	13.8	92	1	YA71_ARCFU	Q29191 archaeoglob
445	4	13.8	43	1	DEPI_PYRAP	P37364 pyrrhocoris	518	4	13.8	93	1	HIS2_MYCLE	Q49786 mycobacteri
446	4	13.8	43	1	PIV6_ADEB2	Q96627 bovine aden	519	4	13.8	93	1	VE7_HPV50	Q80928 human papil
447	4	13.8	50	1	INS_PROGU	P01331 proechimys	520	4	13.8	93	1	Y738_ARCFU	O29520 archaeoglob
448	4	13.8	50	1	RS14_METTH	O26125 methanobact	521	4	13.8	93	1	YQ32_MYCTU	P71931 mycobacteri
449	4	13.8	52	1	GLP_CANFA	P02727 canis famil	522	4	13.8	94	1	NAG7_HUMAN	Q9Y6c7 homo sapien
450	4	13.8	53	1	LEC1_LATOC	P12306 lathyrus oc	523	4	13.8	95	1	MYLE_HUMAN	O95424 homo sapien
451	4	13.8	53	1	LECA_LATAR	P07442 lathyrus ar	524	4	13.8	95	1	MYLE_MOUSE	Q9waq7 mus musculu
452	4	13.8	53	1	RS14_METVA	P14041 methanococc	525	4	13.8	95	1	U119_CANFA	O19177 canis famil
453	4	13.8	55	1	Y072_NFVOP	O10325 orgyia pseu	526	4	13.8	95	1	YWS6_CAEEL	Q10942 caenorhabdi
454	4	13.8	57	1	IVBC_NAJNA	P19859 najia najia	527	4	13.8	96	1	IFIC_TOBAC	P12136 nicotiana t
455	4	13.8	57	1	IVBK_DENAN	P00982 dendroaspis	528	4	13.8	96	1	IMMK_ECOLI	Q47503 escherichia
456	4	13.8	57	1	IVBK_DENPO	P00981 dendroaspis	529	4	13.8	96	1	RR6_ODOSI	P49494 odontella s
457	4	13.8	57	1	SMS2_P1AME	P81879 piaractus m	530	4	13.8	96	1	TRAA_ECOLI	P27187 escherichia
458	4	13.8	59	1	IVBE_DENPO	P00984 dendroaspis	531	4	13.8	97	1	YAJC_HAEIN	P44592 haemophilus
459	4	13.8	60	1	HM09_CAEEL	P56407 caenorhabdi	532	4	13.8	98	1	ELJA_PHYDR	P35696 phytophthor
460	4	13.8	60	1	HMD2_CHICK	Q02814 gallus gall	533	4	13.8	98	1	S113_BOVIN	P79342 bos taurus
461	4	13.8	64	1	BCNB_CLOPE	P15936 clostridium	534	4	13.8	98	1	S113_HUMAN	Q99584 homo sapien
462	4	13.8	64	1	YDP6_SCHPO	O14211 schizosacch	535	4	13.8	98	1	S113_MOUSE	P97352 mus musculu
463	4	13.8	66	1	RS14_TOBAC	P93377 nicotiana t	536	4	13.8	98	1	Y095_MYCLE	Q10891 mycobacteri
464	4	13.8	67	1	SAS2_SPOUR	P52969 sporosarcin	537	4	13.8	99	1	CF10_MYCLE	O33084 mycobacteri
465	4	13.8	69	1	ANX2_PIG	P19620 sus scrofa	538	4	13.8	99	1	PETM_CHLRE	Q42496 chlamydomon
466	4	13.8	69	1	SAS1_SPOUR	P52968 sporosarcin	539	4	13.8	99	1	Y142_UREPA	Q9pr02 ureaplasma
467	4	13.8	70	1	RL28_THEMEA	Q9wy96 thermotoga	540	4	13.8	100	1	DEF6_HUMAN	Q01524 homo sapien
468	4	13.8	70	1	SAS1_BACST	P06552 bacillus st	541	4	13.8	100	1	UCP9_EUGGR	P43266 euglena gra
469	4	13.8	71	1	RR18_MESVI	Q9mup6 mesostigma	542	4	13.8	100	1	WN15_MOUSE	O35468 mus musculu
470	4	13.8	71	1	SAS1_THETP	P06553 thermoactin	543	4	13.8	101	1	CVN_NOSEL	P81180 noscoc ell
471	4	13.8	71	1	VLYS_BPFR	P19903 bacterioph	544	4	13.8	101	1	CY02_CATRO	P25923 catharanthu

545	4	13.8	101	1	RBPA_SYN3	Q57014	synechocyst	618	4	13.8	114	1	Y741_CHLTR	O84746	chlamydia t
546	4	13.8	101	1	VE7_HPV21	P50779	human papil	619	4	13.8	115	1	NU3M_FELCA	P48912	felis silve
547	4	13.8	102	1	NODM_RHILT	Q52846	rhizobium l	620	4	13.8	115	1	PARY_TRYBB	P14043	trypanosoma
548	4	13.8	102	1	RS10_LACIA	Q9cdw1	rhizococcus	621	4	13.8	115	1	PTHY_HUMAN	P01270	homo sapien
549	4	13.8	102	1	VE7_HPV19	P36822	human papil	622	4	13.8	115	1	PTH_MACFA	Q9xt35	macaca fasc
550	4	13.8	102	1	VE7_HPV20	P50778	human papil	623	4	13.8	115	1	RR20_MESVI	Q9mup7	mesostigma
551	4	13.8	102	1	VE7_HPV25	P36823	human papil	624	4	13.8	115	1	SMS1_PROAN	Q9w7f0	protopterus
552	4	13.8	102	1	VE7_HPV36	P50811	human papil	625	4	13.8	115	1	SMS1_RANRI	P87384	rana ridibu
553	4	13.8	102	1	Y224_METH	O26326	methanobact	626	4	13.8	115	1	THIO_MYCTU	P52229	mycobacteri
554	4	13.8	102	1	Y266_BORBU	Q44754	borrelia bu	627	4	13.8	115	1	WNT2_STRPU	P28095	strongyloce
555	4	13.8	102	1	Y182_ECOLI	P08410	escherichia	628	4	13.8	115	1	Y440_ARCFU	P29809	archaeoglob
556	4	13.8	103	1	CBIN_RHOCA	P08104	rhodobacter	629	4	13.8	116	1	ANFG_RHORU	O68940	rhodospiril
557	4	13.8	103	1	REV_HV2BE	P18093	human immun	630	4	13.8	116	1	CN4D_MOUSE	Q01063	mus musculu
558	4	13.8	103	1	REV_HV2D1	P17754	human immun	631	4	13.8	116	1	FDXN_ANASP	P12415	anabaena sp
559	4	13.8	103	1	REV_HV2G1	P18039	human immun	632	4	13.8	116	1	NU3M_ASTPE	P11991	asterina pe
560	4	13.8	103	1	RLAI_POLPE	P27464	polyorchis	633	4	13.8	116	1	NU3M_SQUAC	Q92247	squalus aca
561	4	13.8	103	1	RR10_ODOSI	P49498	odontella s	634	4	13.8	116	1	RL20_MYCFE	Q05427	mycoplasma
562	4	13.8	103	1	VE7_HPV05	P06932	human papil	635	4	13.8	116	1	SMS_CHICK	P33094	gallus gall
563	4	13.8	103	1	VE7_HPV08	P06430	human papil	636	4	13.8	116	1	THIH_PAGES	Q96419	agropyrum e
564	4	13.8	103	1	VE7_HPV12	P36819	human papil	637	4	13.8	117	1	RPBY_DROME	O9vje4	drosophila
565	4	13.8	103	1	VE7_HPV17	P22423	human papil	638	4	13.8	117	1	V13K_TRVPL	P33776	tobacco rat
566	4	13.8	103	1	VE7_HPV5B	P26559	human papil	639	4	13.8	117	1	YFB9_YEAST	P43576	saccharomyc
567	4	13.8	104	1	METJ_HAEIN	P44618	haemophilus	640	4	13.8	117	1	YJEU_ECOLI	P39278	escherichia
568	4	13.8	104	1	RECA_CLOBE	P42441	clostridium	641	4	13.8	118	1	DSR4_HUMAN	P56555	homo sapien
569	4	13.8	104	1	RECA_STRSL	P49987	streptococc	642	4	13.8	118	1	RL20_THEMA	Q9x1s8	thermotoga
570	4	13.8	105	1	SMS2_ICTPU	P01172	ictalurus p	643	4	13.8	118	1	VAG3_HUMAN	O961b4	homo sapien
571	4	13.8	105	1	VE7_HPV66	Q80956	human papil	644	4	13.8	118	1	YCE2_YEAST	P25572	saccharomyc
572	4	13.8	105	1	YJN5_YEAST	P47012	saccharomyc	645	4	13.8	120	1	H2B_MARGE	O92b77	mycoplasma
573	4	13.8	106	1	ILB2_CAEL	Q9627	caenorhabdi	646	4	13.8	120	1	R26A_SCHPO	Q9ut56	schizosacch
574	4	13.8	106	1	THIO_SYN3	P52331	synechocyst	647	4	13.8	119	1	ID3_HUMAN	Q02535	homo sapien
575	4	13.8	106	1	Y857_TREPA	O83829	treponema p	648	4	13.8	119	1	R26B_SCHPO	Q9utg4	schizosacch
576	4	13.8	107	1	DPD4_HUMAN	Q9hcu8	homo sapien	649	4	13.8	119	1	RBFA_IACLA	Q9chg3	lactococcc
577	4	13.8	107	1	DPD4_MOUSE	O9cpw8	mus musculu	650	4	13.8	119	1	Y00E_BPT4	P39415	bacterioph
578	4	13.8	107	1	MREC_ECOLI	P13657	escherichia	651	4	13.8	120	1	H2B_MARGE	O92b77	mycoplasma
579	4	13.8	107	1	RLAI_CHLRE	P29763	chlamydomon	652	4	13.8	120	1	R26A_SCHPO	Q9ut56	schizosacch
580	4	13.8	107	1	RT10_PORNI	P46742	prototheca	653	4	13.8	120	1	RBFA_BUCAL	P57457	buchnera ap
581	4	13.8	107	1	THI2_CROME	P07887	corynebacte	654	4	13.8	120	1	RL20_UREPA	Q9pqir2	ureaplasma
582	4	13.8	107	1	THIO_HAEIN	P43785	haemophilus	655	4	13.8	120	1	RT13_MARPO	P26872	marchantia
583	4	13.8	108	1	HEMA_IARI5	P03450	influenza a	656	4	13.8	120	1	YJY8_YEAST	P47091	saccharomyc
584	4	13.8	108	1	YJAY_YEAST	P47078	saccharomyc	657	4	13.8	121	1	ANRE_HUMAN	P15267	homo sapien
585	4	13.8	109	1	C550_NITWI	P00085	nitrobacter	658	4	13.8	121	1	INSC_ECOLI	P13776	escherichia
586	4	13.8	109	1	RLAI_SCHPO	P17476	schizosacch	659	4	13.8	121	1	SMS1_LOPAM	P01169	lophius ame
587	4	13.8	109	1	YDHA_ECOLI	P28224	escherichia	660	4	13.8	121	1	SRB6_YEAST	P2570	saccharomyc
588	4	13.8	110	1	RLAI_ALTAL	P49148	alterharia	661	4	13.8	122	1	PA21_BOTAS	P20474	bothrops as
589	4	13.8	110	1	RLA3_SCHPO	P17477	schizosacch	662	4	13.8	122	1	VE4_PAPVD	P06925	deer papill
590	4	13.8	110	1	TNG2_HUMAN	P56847	homo sapien	663	4	13.8	123	1	AZUP_METEX	P04171	methyllobact
591	4	13.8	110	1	YKZ1_CAEL	P34328	caenorhabdi	664	4	13.8	123	1	CASK_CAPCA	Q95146	capreolus c
592	4	13.8	110	1	YF12_RTBVP	P27499	rice tungro	665	4	13.8	123	1	CDX1_RAT	Q05095	rattus norv
593	4	13.8	111	1	CARS_MYXXA	Q06911	myxococcus	666	4	13.8	123	1	DHSD_THEAC	Q08717	thermoplasm
594	4	13.8	111	1	H2B_LEIEN	P27893	leishmania	667	4	13.8	123	1	RPBY_SCHPO	P87123	schizosacch
595	4	13.8	111	1	SINR_BACLI	P22753	bacillus li	668	4	13.8	123	1	RS12_BORBU	O51348	borrelia bu
596	4	13.8	111	1	SINR_BACSU	P06533	bacillus su	669	4	13.8	123	1	RS12_CHLPN	O92800	chlamydia p
597	4	13.8	111	1	YR23_CAEL	Q09340	caenorhabdi	670	4	13.8	123	1	RS12_CHLTR	O84446	chlamydia t
598	4	13.8	112	1	ANFG_AZONA	O68946	azomonas ma	671	4	13.8	123	1	RS12_HAEIN	P44412	haemophilus
599	4	13.8	112	1	PERX_PSEPU	P23103	pseudomonas	672	4	13.8	123	1	RS12_NEIMA	Q9jgr6	neisseria m
600	4	13.8	112	1	GLNB_RHIET	O54053	rhizobium e	673	4	13.8	123	1	RS12_PSEAE	Q91hwd0	pseudomonas
601	4	13.8	112	1	MPG1_MAGGR	P52751	magnaporthe	674	4	13.8	124	1	APC4_MOUSE	P61268	mus musculu
602	4	13.8	112	1	PFDL_SCHPO	O14334	schizosacch	675	4	13.8	124	1	RL20_MYCGE	P47440	mycoplasma
603	4	13.8	112	1	Y474_METJA	Q58474	methanococc	676	4	13.8	124	1	RS12_HAEDU	Q956d2	haemophilus
604	4	13.8	113	1	GLNB_CYACA	Q9cm37	cyanidium c	677	4	13.8	124	1	RS12_LEPBI	P20820	leptospira
605	4	13.8	113	1	MEAL_PIG	Q95313	sus scrofa	678	4	13.8	124	1	RS12_PASMU	Q9c186	pasteurella
606	4	13.8	113	1	SUI1_ARATH	P41568	arabidopsis	679	4	13.8	124	1	RS12_THICU	O50563	thiobacillu
607	4	13.8	113	1	SUI1_BRAOL	Q9sqf4	brassica ol	680	4	13.8	124	1	RS12_TREPA	O83271	treponema p
608	4	13.8	114	1	HYPA_CAMJE	Q9php0	campylobact	681	4	13.8	124	1	RS12_VIBCH	O9kuz9	vibrio chol
609	4	13.8	114	1	PARA_TRYBB	P18764	trypanosoma	682	4	13.8	124	1	RS12_XYLF	Q9pa88	xytella fas
610	4	13.8	114	1	RLAI_CHICK	P18660	gallus gall	683	4	13.8	125	1	VSIS_REOVJ	P32868	reovirus (t
611	4	13.8	114	1	RLAI_HUMAN	P05386	homo sapien	684	4	13.8	126	1	PROF_DROME	P25843	drosophila
612	4	13.8	114	1	RLAI_MOUSE	P47955	mus musculu	685	4	13.8	126	1	SYGB_NEIGO	Q50945	neisseria g
613	4	13.8	114	1	RLAI_RAT	P19944	rattus norv	686	4	13.8	126	1	YJGG_ECOLI	P27843	escherichia
614	4	13.8	114	1	SMS1_ICTPU	P01171	ictalurus p	687	4	13.8	127	1	GR14_NEOCA	Q25540	neospora ca
615	4	13.8	114	1	SMSA_CARAU	Q9ygh5	carassius a	688	4	13.8	127	1	KP4T_DMV4	Q90121	ustilago ma
616	4	13.8	114	1	VG40_BPT4	P17171	bacterioph	689	4	13.8	127	1	RL20_MYCPN	P78023	mycoplasma
617	4	13.8	114	1	Y117_CHLMU	Q9pl12	chlamydia m	690	4	13.8	127	1	RS12_THEMA	Q9x1j3	thermotoga

691	4	13.8	128	1	ANF_CAVPO	P27596	cavia porce	764	4	13.8	140	1	ATPE_VIBCH	O9knh6
692	4	13.8	128	1	B2_TENMO	Q27018	tenebrio mo	765	4	13.8	140	1	DUT_SCHPO	O9p6d5
693	4	13.8	128	1	CRB2_HALN1	O9hnl1	halobacteri	766	4	13.8	140	1	MOPB_ERWCA	O34199
694	4	13.8	128	1	E314_ADE02	P03249	human adeno	767	4	13.8	140	1	YB38_YEAST	O38302
695	4	13.8	128	1	E314_ADE05	P04493	human adeno	768	4	13.8	140	1	YCPA_SYNPY	O38302
696	4	13.8	128	1	RL7_THEMEA	P29396	thermotoga	769	4	13.8	140	1	YI47_MYCTU	O28186
697	4	13.8	128	1	RS12_CAMJE	O9p118	campylobact	770	4	13.8	141	1	DUT_CHVP1	P95162
698	4	13.8	128	1	Y192_RICPR	O9zdx2	rickettsia	771	4	13.8	141	1	HBA3_XENTR	O41033
699	4	13.8	128	1	YD8A_HAEIN	O86237	haemophilus	772	4	13.8	141	1	HBA3_XENTR	P02006
700	4	13.8	129	1	ANFC_RANCA	P20968	rana catesb	773	4	13.8	141	1	V223_FOWPV	O91512
701	4	13.8	129	1	PARB_TRYBB	P09791	trypanosoma	774	4	13.8	141	1	YC59_LISIN	O926a6
702	4	13.8	129	1	PURL_METTM	P26500	methanobact	775	4	13.8	141	1	YFAO_ECOLI	P52006
703	4	13.8	129	1	RECA_BACTN	O45791	bacteroides	776	4	13.8	141	1	YFAO_ECOLI	P21498
704	4	13.8	130	1	BI_TENMO	Q27017	tenebrio mo	777	4	13.8	142	1	HBA_TARGR	P02014
705	4	13.8	130	1	ITP_SCHGR	Q26491	schistocerc	778	4	13.8	142	1	RECA_STRCT	P41055
706	4	13.8	130	1	OREX_MOUSE	O55241	mus musculus	779	4	13.8	143	1	MERC_THIFE	P22905
707	4	13.8	130	1	OREX_RAT	O55232	rattus norv	780	4	13.8	143	1	PARL_TRYBB	P08459
708	4	13.8	130	1	RS22_SCHPO	O14469	schizosacch	781	4	13.8	143	1	R19B_SCHPO	P79016
709	4	13.8	130	1	SZ05_RAT	P97885	rattus norv	782	4	13.8	143	1	ROSR_RHET	O52741
710	4	13.8	130	1	YT64_CAEEL	Q11080	caenorhabdi	783	4	13.8	143	1	YCX1_SOYBN	P09363
711	4	13.8	131	1	ATPQ_CAEEL	Q18803	caenorhabdi	784	4	13.8	143	1	YIFN_HAEIN	P43936
712	4	13.8	131	1	ATPQ_CAEEL	P09021	caenorhabdi	785	4	13.8	143	1	YL66_MYCTU	O06211
713	4	13.8	131	1	CALN_CHICK	P05419	gallus gall	786	4	13.8	144	1	DUT_VACCV	P17374
714	4	13.8	131	1	FABI_XENLA	Q91775	xenopus lae	787	4	13.8	144	1	MARR_SALTY	O56069
715	4	13.8	131	1	FER2_METJA	O57563	methanococc	788	4	13.8	144	1	R19A_SCHPO	P58234
716	4	13.8	131	1	FRDC_ECOLI	P03805	escherichia	789	4	13.8	144	1	RL15_BUCAK	P46185
717	4	13.8	131	1	MCRD_METFE	P12974	methanother	790	4	13.8	144	1	RS14_TRYBB	P19800
718	4	13.8	131	1	PPK_CHLTE	O68984	chlorobium	791	4	13.8	144	1	RS19_HUMAN	P39019
719	4	13.8	131	1	SPEK_THEAC	Q9hly0	thermoplasma	792	4	13.8	144	1	RS19_RAT	P17074
720	4	13.8	131	1	TRY1_ECOLI	P06627	escherichia	793	4	13.8	145	1	ANF_RANCA	P18909
721	4	13.8	131	1	VMU_LAMB2	P03732	bacterioph	794	4	13.8	145	1	DUT_FOWPV	O91545
722	4	13.8	132	1	VE4_HPV2A	P25483	human papil	795	4	13.8	145	1	OM20_HUMAN	O15388
723	4	13.8	133	1	R24B_SCHPO	O59865	schizosacch	796	4	13.8	145	1	OM20_RAT	O62760
724	4	13.8	133	1	RSBT_BACSU	P42411	bacillus su	797	4	13.8	145	1	PARC_TRYBB	O60084
725	4	13.8	133	1	Y496_HELPJ	Q9zlx8	helicobacte	798	4	13.8	145	1	RS19_MTXGL	Q9y0h3
726	4	13.8	133	1	Y496_HELPJ	P94842	helicobacte	799	4	13.8	145	1	YB80_METJA	O58580
727	4	13.8	133	1	YPI5_METTM	P14934	methanobact	800	4	13.8	146	1	ES16_ADE03	P11320
728	4	13.8	134	1	ATFB_HUMAN	P16860	homo sapien	801	4	13.8	146	1	FLGN_PROMI	P96975
729	4	13.8	134	1	ITPL_SCHGR	Q26492	schistocerc	802	4	13.8	146	1	FUR_HAEIN	P44561
730	4	13.8	134	1	LEGL_RAT	P11762	rattus norv	803	4	13.8	146	1	HBG_GADMO	O13077
731	4	13.8	134	1	RR9_GUTH	P19459	guillardia	804	4	13.8	146	1	RS19_CAEEL	O18650
732	4	13.8	134	1	RS14_TORRU	Q9xek6	tortula rur	805	4	13.8	146	1	VE6_HPV61	O80948
733	4	13.8	134	1	SYW_ENCCU	O96771	encephalito	806	4	13.8	146	1	YG5G_YEAST	P53321
734	4	13.8	134	1	YF60_ARCFU	O28712	archaeoglob	807	4	13.8	147	1	CALM_FAGSY	Q39752
735	4	13.8	135	1	H34_CAIMO	P06902	cairina mos	808	4	13.8	147	1	CALM_KLULA	O60041
736	4	13.8	135	1	MERR_STAAU	P22874	staphylococ	809	4	13.8	147	1	CALM_YEAST	P06787
737	4	13.8	135	1	RS12_HELPY	P56019	helicobacte	810	4	13.8	147	1	DUT_VACCC	P21035
738	4	13.8	135	1	RS12_THETH	P17293	thermus aqu	811	4	13.8	147	1	DUT_VARY	P33826
739	4	13.8	135	1	RS19_PIG	Q29308	sus scrofa	812	4	13.8	147	1	DUT_YEAST	P33317
740	4	13.8	136	1	CALI_ARATH	P25854	arabidopsis	813	4	13.8	147	1	FLGN_YEREN	O56854
741	4	13.8	136	1	VNS1_ORSVW	O65703	ovine respi	814	4	13.8	147	1	LECE_ANTCR	P06027
742	4	13.8	136	1	YNN6_YEAST	P53864	saccharomyc	815	4	13.8	147	1	YPMF_ECOLI	P07116
743	4	13.8	137	1	RS12_BACSU	P21472	bacillus su	816	4	13.8	148	1	AZUR_METFL	O50400
744	4	13.8	137	1	RS12_LACIA	Q9cdf9	lactococcus	817	4	13.8	148	1	CALI_PETHY	P27162
745	4	13.8	137	1	RS12_STAAM	P48942	staphylococ	818	4	13.8	148	1	CAL2_ARATH	P25089
746	4	13.8	137	1	RS12_STRGN	Q9f0r4	streptococc	819	4	13.8	148	1	CAL2_PETHY	P27163
747	4	13.8	137	1	RS12_STRPN	P30891	streptococc	820	4	13.8	148	1	CAL4_ARATH	O03510
748	4	13.8	137	1	RS12_STRPY	P58172	streptococc	821	4	13.8	148	1	CAL6_ARATH	O03509
749	4	13.8	137	1	RT16_HUMAN	Q9y3d3	homo sapien	822	4	13.8	148	1	CALM_ACHKL	P15094
750	4	13.8	137	1	YUXK_BACSU	P40761	bacillus su	823	4	13.8	148	1	CALM_CAPAN	P93087
751	4	13.8	138	1	CALL_ARBPJ	P05932	arabacia pun	824	4	13.8	148	1	CALM_DROME	P07181
752	4	13.8	138	1	END5_BPT4	P04418	bacterioph	825	4	13.8	148	1	CALM_ELEEL	P02594
753	4	13.8	138	1	SCHB_STRHA	Q05362	streptomyce	826	4	13.8	148	1	CALM_EMENI	P19533
754	4	13.8	138	1	Y922_METJA	O58332	methanococc	827	4	13.8	148	1	CALM_HELAN	P93171
755	4	13.8	139	1	ANGL_CHICK	P27043	gallus gall	828	4	13.8	148	1	CALM_HORVU	P13565
756	4	13.8	139	1	LYC_HYACE	P05105	hyalophora	829	4	13.8	148	1	CALM_HUMAN	P02593
757	4	13.8	139	1	RS12_BACHD	Q9z319	bacillus ha	830	4	13.8	148	1	CALM_MAGGR	P27161
758	4	13.8	139	1	RS12_BACST	P09901	bacillus st	831	4	13.8	148	1	CALM_LYCES	Q9uwf0
759	4	13.8	139	1	RS14_SCHPO	O14150	schizosacch	832	4	13.8	148	1	CALM_MAIZE	P41040
760	4	13.8	139	1	RSFR_CHICK	P30374	gallus gall	833	4	13.8	148	1	CALM_MALDO	P48976
761	4	13.8	139	1	UMUD_ECOLI	P04153	escherichia	834	4	13.8	148	1	CALM_MEDSA	P17928
762	4	13.8	139	1	YRC2_CAEEL	O10041	caenorhabdi	835	4	13.8	148	1	CALM_METSE	P02596
763	4	13.8	140	1	ATPE_VIBAL	P12988	vibrio algi	836	4	13.8	148	1	CALM_NEUCR	O02052

837	4	13.8	148	1	CALM_ORYSA	P29612	oryza sativ	910	4	13.8	158	1	FLAV_BACSU	O34737	bacillus su
838	4	13.8	148	1	CALM_PATSP	P02595	patinopecte	911	4	13.8	158	1	RK12_ORYSA	O2386	oryza sativ
839	4	13.8	148	1	CALM_PHYTN	P27165	phytophthor	912	4	13.8	158	1	RL15_AERPE	O9vf98	aeropyrum p
840	4	13.8	148	1	CALM_PLECO	P11120	pleurotus c	913	4	13.8	158	1	UBCI_MSAU	O0181	mesocricetu
841	4	13.8	148	1	CALM_PYUSP	P11121	pyuridae sp	914	4	13.8	158	1	VE6_HPV16	P03126	human papil
842	4	13.8	148	1	CALM_SOLUTU	P13868	solanum tub	915	4	13.8	158	1	YBUN_ECOLI	P75815	escherichia
843	4	13.8	148	1	CALM_SPIOL	P04353	spinacia ol	916	4	13.8	158	1	YC76_MYCTU	Q11043	mycobacteri
844	4	13.8	148	1	CALM_STIJA	P21251	stichopus j	917	4	13.8	158	1	YCBM_BACSU	P42245	bacillus su
845	4	13.8	148	1	CLM_PLEOS	O94739	pleurotus o	918	4	13.8	158	1	YMI9_BRANA	P42245	bacillus su
846	4	13.8	148	1	GRP_HUMAN	P07492	homo sapien	919	4	13.8	158	1	YMI9_BRANA	O01552	brassica na
847	4	13.8	148	1	PBPI_DRONE	P54191	drosofila	920	4	13.8	158	1	DUT_CANAL	O06564	raphanus sa
848	4	13.8	148	1	RI96_ASCSU	P24494	ascaris suu	921	4	13.8	159	1	DUT_ORFN2	P43058	candida alb
849	4	13.8	148	1	RS19_ENTHI	O15631	entamoeba h	922	4	13.8	159	1	EUTP_SALTY	P14597	orf virus (
850	4	13.8	148	1	UBC4_LYCES	P35135	lycopersico	923	4	13.8	159	1	IF51_NICPL	Q94fv6	salmonella
851	4	13.8	149	1	ANF_CANFA	P07499	canis famil	924	4	13.8	159	1	IF52_NICPL	P24921	nicotiana p
852	4	13.8	149	1	ARGR_BACST	O31408	bacillus st	925	4	13.8	159	1	IF54_SOLUTU	P24922	nicotiana p
853	4	13.8	149	1	CALM_CANAL	P23286	candida alb	926	4	13.8	159	1	IF55_SOLUTU	P56336	solanum tub
854	4	13.8	149	1	CALM_WHEAT	P04464	tritium ae	927	4	13.8	159	1	VE6_HPV27	P56337	solanum tub
855	4	13.8	149	1	RI41_MAIZE	P19950	zea mays (m	928	4	13.8	159	1	Y207_MYCPN	P36808	human papil
856	4	13.8	149	1	RS19_MYAAR	Q94613	mya arenari	929	4	13.8	159	1	YMI9_HELAN	P75349	mycoplasma
857	4	13.8	149	1	SSB_BORBU	O51141	borrelia bu	930	4	13.8	159	1	YMI9_OENBE	P41248	helianthus
858	4	13.8	149	1	TCMJ_STRGA	P16558	streptomyce	931	4	13.8	160	1	BIK_HUMAN	P08746	oenothera b
859	4	13.8	149	1	YB07_METJA	Q58507	methanococc	932	4	13.8	160	1	IF51_SOLUTU	Q13323	homo sapien
860	4	13.8	150	1	ANF_PIG	P24259	sus scrofa	933	4	13.8	160	1	IF53_SOLUTU	P56333	solanum tub
861	4	13.8	150	1	RI42_MAIZE	P19951	zea mays (m	934	4	13.8	160	1	IF5A_MAIZE	P56335	solanum tub
862	4	13.8	150	1	RI9S_ASCSU	P39698	ascaris suu	935	4	13.8	160	1	MOAC_HAEIN	P80639	zea mays (m
863	4	13.8	151	1	DUT_ECOLI	P06968	escherichia	936	4	13.8	161	1	ASFL_HELAN	P45310	haemophilus
864	4	13.8	151	1	LE14_GOSHI	P46518	gossypium h	937	4	13.8	161	1	CALL_CAEEL	P23357	helianthus
865	4	13.8	151	1	RNB_HSV2H	P89479	herpes simp	938	4	13.8	161	1	GCS3_AQUAE	P04630	caenorhabdi
866	4	13.8	152	1	ACPM_DRONE	Q94519	drosofila	939	4	13.8	161	1	PTI5_LYCES	O67080	aquifex aeo
867	4	13.8	152	1	ANF_BOVIN	P07501	bos taurus	940	4	13.8	161	1	YC76_MYCLE	O04681	lycopersico
868	4	13.8	152	1	LE14_SOYBN	P46519	glycine max	941	4	13.8	161	1	MRED_HAEIN	Q9cc85	mycobacteri
869	4	13.8	152	1	SMN1_SCHPO	Q09808	schizosacch	942	4	13.8	162	1	PTSN_KLEPN	P44476	haemophilus
870	4	13.8	152	1	ULC4_CHMVA	P16742	human cytom	943	4	13.8	162	1	RL10_BORBU	P17162	klebsiella
871	4	13.8	152	1	Y396_MYCCE	P47636	mycoplasma	944	4	13.8	162	1	SODC_LEGPN	O51352	borrelia bu
872	4	13.8	152	1	Y554_METJA	Q57974	methanococc	945	4	13.8	162	1	Y013_METJA	P53637	legionella
873	4	13.8	152	1	YF63_METJA	Q58958	methanococc	946	4	13.8	162	1	Y096_GVCL	Q60321	methanococc
874	4	13.8	153	1	ANF_HORSE	P27104	equus cabal	947	4	13.8	163	1	DUT_ADEG8	P41728	cryptophleb
875	4	13.8	153	1	ANF_HUMAN	P01160	homo sapien	948	4	13.8	163	1	HL_TETHH	Q9yy50	avian adeno
876	4	13.8	153	1	ANF_RABIT	P07500	oryctolagus	949	4	13.8	163	1	PTSN_ECOLI	P10156	tetrahymena
877	4	13.8	153	1	HEMA_MUMPS	P33480	mumps virus	950	4	13.8	163	1	V17_BPT3	P31222	escherichia
878	4	13.8	153	1	MLR_PATSP	P02613	patinopecte	951	4	13.8	163	1	V18K_MLVAB	P07719	bacterioph
879	4	13.8	153	1	MYG_CTEGU	P20856	ctenodactyl	952	4	13.8	163	1	CD32_MOUSE	P03400	abelson mur
880	4	13.8	153	1	MYG_LAGMA	P04250	lagostomus	953	4	13.8	164	1	DNAT_BUCAL	P24161	mus musculus
881	4	13.8	153	1	MYG_PROGU	P04249	proechimys	954	4	13.8	164	1	EBSC_ENTFA	P57135	buchnera ap
882	4	13.8	153	1	RS14_CHLRE	P46295	chlamydomon	955	4	13.8	164	1	GSSH_CHICK	P36922	enterococu
883	4	13.8	153	1	VE6_HPV2A	P25484	human papil	956	4	13.8	164	1	MCRW_METJA	P11183	gallus gall
884	4	13.8	153	1	Y992_SYNF3	P74516	synechocyst	957	4	13.8	164	1	DEST_HUMAN	Q58253	methanococ
885	4	13.8	153	1	YPH1_SYNP2	P32039	synechococc	958	4	13.8	165	1	PHAZ_SYNP2	P18282	homo sapien
886	4	13.8	154	1	DUT_MYCTU	O07199	mycobacteri	959	4	13.8	165	1	POC2_JUNOX	P37720	synechococc
887	4	13.8	154	1	IL2_PIG	P26891	sus scrofa	960	4	13.8	165	1	RL10_BUCAL	O64943	juniperus o
888	4	13.8	154	1	IL7_RAT	P56478	rattus norv	961	4	13.8	165	1	Y400_SYNP2	P57148	buchnera ap
889	4	13.8	154	1	KAMB_STRTN	P25920	streptomyce	962	4	13.8	166	1	VB10_VACCV	Q54751	synechococc
890	4	13.8	154	1	NER2_EUPOC	O15825	euploetes oc	963	4	13.8	167	1	ARGR_MYCLE	P21006	vaccinia vi
891	4	13.8	154	1	OM22_NEUCR	Q07335	neurospora	964	4	13.8	167	1	F6B1_ECOLI	P57992	mycobacteri
892	4	13.8	154	1	PF25_SCHPO	O94307	schizosacch	965	4	13.8	167	1	F6B2_ECOLI	P53510	escherichia
893	4	13.8	154	1	PYRI_BUCAL	P57451	buchnera ap	966	4	13.8	167	1	TCTP_LUMRU	P53511	escherichia
894	4	13.8	154	1	SODC_BRARE	O73872	brachydanio	967	4	13.8	167	1	Y199_MYCCE	O18477	lumbicus r
895	4	13.8	154	1	X_HPVIL	P12912	hepatitis b	968	4	13.8	168	1	CYNS_ARATH	P47441	mycoplasma
896	4	13.8	154	1	YD88_SYNF3	P74148	synechocyst	969	4	13.8	169	1	DUT_LYCES	O22683	arabidopsis
897	4	13.8	155	1	FER6_MAIZE	P94044	zea mays (m	970	4	13.8	169	1	RL10_RICCN	P32518	lycopersico
898	4	13.8	155	1	IL17_HUMAN	Q16552	homo sapien	971	4	13.8	169	1	SULA_SALTY	Q92388	rickettsia
899	4	13.8	155	1	PA2X_HUMAN	O15496	homo sapien	972	4	13.8	169	1	YAJO_ECOLI	P08847	salmonella
900	4	13.8	155	1	RS19_DRONE	P39018	drosofila	973	4	13.8	169	1	YSAL_PNECA	P77482	escherichia
901	4	13.8	156	1	IF51_YEAST	P19211	saccharomyc	974	4	13.8	170	1	NEUT_HUMAN	O01688	pneumocysti
902	4	13.8	156	1	MLR_AEQIR	P13543	aequipecten	975	4	13.8	170	1	YCSB_BACSU	P30990	homo sapien
903	4	13.8	156	1	MLR_CHLNI	P05963	chlamys nip	976	4	13.8	171	1	CP18_RABIT	P42961	bacillus su
904	4	13.8	156	1	PRO2_ORYSA	P17048	oryza sativ	977	4	13.8	171	1	NSG2_MOUSE	P25230	oryctolagus
905	4	13.8	156	1	YMI9_WHEAT	P43650	tritium ae	978	4	13.8	171	1	R9B7_YEAST	P47759	mus musculus
906	4	13.8	157	1	WH16_STRCO	P23157	streptomyce	979	4	13.8	171	1	TRAV_ECOLI	P34087	saccharomyc
907	4	13.8	157	1	Y118_NPVAC	P41671	autographa	980	4	13.8	171	1	YAPD_RHISN	P41069	escherichia
908	4	13.8	157	1	Y188_RICPR	Q92dx6	rickettsia	981	4	13.8	171	1	YTHI_SALTY	P55613	rhizobium s
909	4	13.8	157	1	YQAA_HAEIN	P44005	haemophilus	982	4	13.8	171	1	YSNB_BACSU	P37130	salmonella
														P94559	bacillus su

983	4	13.8	172	1	INT3_SHEEP	P56832	ovis aries	1056	4	13.8	185	1	VP4_SBMV	P21406	southern be
984	4	13.8	172	1	MEAL_HUMAN	Q16626	homo sapien	1057	4	13.8	185	1	YACH_BACSU	P37569	bacillus su
985	4	13.8	172	1	MRED_BACSU	Q01467	bacillus su	1058	4	13.8	185	1	YEP6_YEAST	P40042	saccharomyc
986	4	13.8	172	1	PHE2_PSEA9	Q52451	pseudanabae	1059	4	13.8	186	1	ARL6_HUMAN	Q910f7	homo sapien
987	4	13.8	172	1	PHEB_MASIA	P00313	mastigoclad	1060	4	13.8	186	1	ARL6_MOUSE	O88848	mus musculus
988	4	13.8	172	1	PB03_MYCPN	P75566	mycoplasma	1061	4	13.8	186	1	SODC_NEIMA	P57005	neisseria m
989	4	13.8	173	1	CH19_DRONE	P07186	drosophila	1062	4	13.8	186	1	SODC_NEIMA	Q55623	neisseria m
990	4	13.8	173	1	GLBH_TRICO	P27613	trichostrom	1063	4	13.8	186	1	SODC_PASMU	Q59689	pasteurella
991	4	13.8	173	1	SODC_PHOLE	P00446	photobacter	1064	4	13.8	186	1	UL55_HSV11	P10239	herpes simp
992	4	13.8	174	1	CSF3_SHEEP	Q28746	ovis aries	1065	4	13.8	186	1	UL55_HSV11	P36296	herpes simp
993	4	13.8	174	1	H222_DRONE	P02515	drosophila	1066	4	13.8	186	1	UL55_HSV2H	P28281	herpes simp
994	4	13.8	174	1	MEAL_BOVIN	Q29407	bos taurus	1067	4	13.8	186	1	VMA2_TRTV	P33494	turkey rhin
995	4	13.8	174	1	MEAL_MOUSE	Q64327	mus musculus	1068	4	13.8	186	1	YQGE_HAEIN	P43980	haemophilus
996	4	13.8	174	1	VG17_BPZPA	P08389	bacterioph	1069	4	13.8	187	1	KIP2_HUMAN	O75838	homo sapien
997	4	13.8	175	1	AROK_PASMU	P57925	pasteurella	1070	4	13.8	187	1	KIP2_MOUSE	Q92309	mus musculus
998	4	13.8	175	1	CSF3_CANFA	P35834	canis famil	1071	4	13.8	187	1	MT28_YEAST	P40573	saccharomyc
999	4	13.8	175	1	EPO_CANFA	P33707	canis famil	1072	4	13.8	187	1	SODC_HAEIN	P25841	haemophilus
1000	4	13.8	175	1	HOBH_ECOLI	P36558	escherichia	1073	4	13.8	187	1	SODC_HAEIN	P25842	haemophilus
1001	4	13.8	176	1	C561_YERPE	Q9X6b2	yersinia pe	1074	4	13.8	188	1	ADML_CANFA	O62827	bos taurus
1002	4	13.8	176	1	CYTL_STOHE	P81662	stoichactis	1075	4	13.8	188	1	ADML_CANFA	O77559	canis famil
1003	4	13.8	176	1	ISPF_CHLPN	Q92805	chlamydia p	1076	4	13.8	188	1	ADML_PIG	P53366	sus scrofa
1004	4	13.8	176	1	RL10_THEAC	Q9h1b3	thermoplasm	1077	4	13.8	188	1	COAT_EPMV	P20123	eggplant mo
1005	4	13.8	176	1	RL10_THEVO	P58299	thermoplasm	1078	4	13.8	189	1	COAT_TYMW	P03608	turnip yell
1006	4	13.8	176	1	RWNG_ASPRE	P04389	aspergillus	1079	4	13.8	189	1	COAT_TYMW	P20125	turnip yell
1007	4	13.8	176	1	TRAF_AGRF5	Q44350	agrobacteri	1080	4	13.8	189	1	TBP_AERPE	Q9yat1	aeropyrum p
1008	4	13.8	177	1	BFT_ARATH	Q9fit4	arabidopsis	1081	4	13.8	190	1	EPO_PIG	P49157	sus scrofa
1009	4	13.8	177	1	COP2_BOVIN	P35604	bos taurus	1082	4	13.8	190	1	NUGM_ARATH	Q95748	arabidopsis
1010	4	13.8	177	1	COP2_HUMAN	Q9y3c3	homo sapien	1083	4	13.8	190	1	NUGM_SOLTU	P80261	solanum tub
1011	4	13.8	177	1	CUF2_SCHPO	Q94588	schizosacch	1084	4	13.8	190	1	RNH2_SYNT3	P72657	synecocyst
1012	4	13.8	177	1	IF43_YEAST	P12962	saccharomyc	1085	4	13.8	190	1	RS9_FODAN	P52810	podospora a
1013	4	13.8	177	1	IL19_HUMAN	Q9uhd0	homo sapien	1086	4	13.8	190	1	SODC_ACTPL	P24702	actinobacil
1014	4	13.8	177	1	PB1L_TOBAC	P11670	nicotiana t	1087	4	13.8	190	1	Y516_AERPE	Q9yer5	aeropyrum p
1015	4	13.8	177	1	PSS_BACSU	P39823	bacillus su	1088	4	13.8	191	1	CLD7_RAT	Q92111	rattus norv
1016	4	13.8	177	1	SODC_SALTY	P53636	salmonella	1089	4	13.8	191	1	KTHY_HELPJ	Q921e5	helicobacte
1017	4	13.8	177	1	Y992_METJA	Q58399	methanococc	1090	4	13.8	191	1	KTHY_HELPJ	Q921e5	helicobacte
1018	4	13.8	177	1	YH84_ARCFU	Q28490	archaeoglob	1091	4	13.8	191	1	RPOE_STRPY	P36053	streptococc
1019	4	13.8	178	1	DUT_ADEGI	Q89662	avian adeno	1092	4	13.8	192	1	EPO_FELCA	P33708	felis silve
1020	4	13.8	179	1	AFT_HUMAN	P07741	homo sapien	1093	4	13.8	192	1	GCH2_HELPJ	Q92142	helicobacte
1021	4	13.8	179	1	COO7_RAT	Q63619	rattus norv	1094	4	13.8	192	1	GCH2_HELPJ	Q92142	helicobacte
1022	4	13.8	180	1	AROK_HAEIN	P43880	haemophilus	1095	4	13.8	192	1	NUGM_BETR	Q33994	beta trigyn
1023	4	13.8	180	1	FMF3_ECOLI	P11312	escherichia	1096	4	13.8	192	1	NUGM_BETR	Q33994	beta trigyn
1024	4	13.8	180	1	IL17B_HUMAN	Q9uhf5	homo sapien	1097	4	13.8	192	1	NUGM_BETR	Q33994	beta trigyn
1025	4	13.8	180	1	TAAS_WHEAT	P16347	trititum ae	1098	4	13.8	192	1	NUGM_BETR	Q33994	beta trigyn
1026	4	13.8	180	1	NUOI_BUCAI	P37259	buchnera ap	1099	4	13.8	192	1	PAAD_CHLMU	Q9pkh2	chlamydia m
1027	4	13.8	180	1	ORN_SCHFO	Q94626	schizosacch	1100	4	13.8	192	1	PAAD_CHLMU	Q9pkh2	chlamydia m
1028	4	13.8	180	1	PTTG_HUMAN	P53801	homo sapien	1101	4	13.8	192	1	PHSB_SALTY	P37601	salmonella
1029	4	13.8	180	1	YB27_MYCPN	P75348	mycoplasma	1102	4	13.8	192	1	POLG_HCV1	P27954	hepatitis c
1030	4	13.8	181	1	CAT1_PARTE	Q27177	paramecium	1103	4	13.8	193	1	SODF_BORPE	P37369	borderetella
1031	4	13.8	182	1	CAT2_PARTE	Q27178	paramecium	1104	4	13.8	193	1	FLPA_METNA	O53133	methanosarc
1032	4	13.8	182	1	COX2_RECAM	Q27179	paramecium	1105	4	13.8	193	1	MOBA_RHOSH	P56645	rhodobacter
1033	4	13.8	182	1	IT2_PSOTE	Q21243	reclinomona	1106	4	13.8	194	1	AP22_YEAST	P47064	saccharomyc
1034	4	13.8	182	1	REC6_SCHPO	P25700	psophocarpu	1107	4	13.8	194	1	CSF3_FELCA	O02708	felis silve
1035	4	13.8	183	1	CAL3_PETHY	P40385	schizosacch	1108	4	13.8	194	1	KTHY_METTH	Q27793	methanobact
1036	4	13.8	183	1	CAT3_PARTE	P27164	petunia hyb	1109	4	13.8	194	1	PTH_PSEAE	Q9hvc3	pseudomonas
1037	4	13.8	183	1	DUT_STRCO	Q27178	paramecium	1110	4	13.8	195	1	BETI_ECOLI	P17446	escherichia
1038	4	13.8	183	1	LPOE_MYCLE	O54134	streptomyce	1111	4	13.8	195	1	CSF3_BOVIN	P35833	bos taurus
1039	4	13.8	183	1	UBC6_ARATH	Q9z8m7	mycobacteri	1112	4	13.8	195	1	CSF3_PIG	O2837	sus scrofa
1040	4	13.8	183	1	YB00_MYCPN	P42750	arabidopsis	1113	4	13.8	195	1	HYAD_ECOLI	P19930	escherichia
1041	4	13.8	183	1	YBCL_ECOLI	P75592	mycoplasma	1114	4	13.8	195	1	INT1_SHEEP	P56828	ovis aries
1042	4	13.8	183	1	YJRO_YEAST	P77368	escherichia	1115	4	13.8	195	1	INT2_SHEEP	P56829	ovis aries
1043	4	13.8	184	1	MP20_DRONE	P46993	saccharomyc	1116	4	13.8	195	1	INT4_SHEEP	Q28594	ovis aries
1044	4	13.8	184	1	RK5_MESVI	P14318	drosophila	1117	4	13.8	195	1	INT5_SHEEP	Q28595	ovis aries
1045	4	13.8	184	1	Y180_CHLPN	Q9nuu5	mesostigma	1118	4	13.8	195	1	INT6_SHEEP	Q29429	ovis aries
1046	4	13.8	184	1	Y803_CHLPN	Q9z904	chlamydia p	1119	4	13.8	195	1	INT7_SHEEP	Q08071	ovis aries
1047	4	13.8	184	1	YCDY_HAEIN	P44248	haemophilus	1120	4	13.8	195	1	INT8_SHEEP	Q08072	ovis aries
1048	4	13.8	185	1	ADML_HUMAN	P35318	homo sapien	1121	4	13.8	195	1	INT9_SHEEP	Q08073	ovis aries
1049	4	13.8	185	1	DCRB_ECOLI	P37620	escherichia	1122	4	13.8	195	1	INTA_SHEEP	Q08053	ovis aries
1050	4	13.8	185	1	RL25_CHLPN	Q9plc2	chlamydia m	1123	4	13.8	195	1	INTB_SHEEP	P28169	ovis aries
1051	4	13.8	185	1	RL25_CHLPN	Q926v7	chlamydia p	1124	4	13.8	195	1	INT_CAPHI	P28171	capra hircu
1052	4	13.8	185	1	RL25_CHLTR	O84805	chlamydia t	1125	4	13.8	195	1	INT_OVIMO	P28172	ovibos mosc
1053	4	13.8	185	1	RRF_STRCO	Q46293	clostridium	1126	4	13.8	195	1	KGUA_MYCGA	Q9kx62	mycoplasma
1054	4	13.8	185	1	RRF_STRCO	O86770	streptomyce	1127	4	13.8	195	1	RUBY_CLOPE	P51591	clostridium
1055	4	13.8	185	1	SC6_SCHCO	Q074300	schizophyll	1128	4	13.8	195	1	TSAL_YEAST	P34760	saccharomyc
														Q04120	saccharomyc

1129	4	13.8	196	1	3MGH_BACSU	P94378	bacillus su
1130	4	13.8	196	1	CRBB_HUMAN	P53672	homo sapien
1131	4	13.8	196	1	CRBB_MOUSE	Q91jvl	mus musculus
1132	4	13.8	196	1	GCH2_ECOLI	P25523	escherichia
1133	4	13.8	196	1	HTGA_ECOLI	P28697	escherichia
1134	4	13.8	196	1	RBMS_HUMAN	Q93062	homo sapien
1135	4	13.8	196	1	RBMS_XENLA	Q9ygp5	xenopus lae
1136	4	13.8	196	1	TSAL_CANAL	Q9y7f0	candida alb
1137	4	13.8	196	1	V17_BPT7	P03781	bacterioph
1138	4	13.8	196	1	X349_CHLTR	O84353	chlamydia t
1139	4	13.8	196	1	Y40T_RHLSN	P55605	rhizobium s
1140	4	13.8	196	1	Y628_CHLMW	Q9pk45	chlamydia m
1141	4	13.8	196	1	YF01_MYCPN	P75286	mycoplasma
1142	4	13.8	196	1	YF71_CAEEL	Q09220	caenorhabdi
1143	4	13.8	196	1	YS64_CAEEL	Q09380	caenorhabdi
1144	4	13.8	197	1	RASB_DICDI	P32252	dictyosteli
1145	4	13.8	197	1	RBMS_MOUSE	Q9wv00	mus musculus
1146	4	13.8	197	1	YCF4_CHLRE	O20030	chlamydomon
1147	4	13.8	197	1	YF79_ARCFU	O28693	archaeoglob
1148	4	13.8	198	1	CDNB_CRIGR	Q60439	cricketulus
1149	4	13.8	198	1	TTK_ECOLI	P06969	escherichia
1150	4	13.8	199	1	GRAD_HUMAN	P78560	homo sapien
1151	4	13.8	199	1	GRAD_MOUSE	O88843	mus musculus
1152	4	13.8	199	1	THY_HALN1	Q9hmv4	halobacteri
1153	4	13.8	199	1	MLO3_SCHPO	Q09330	schizosacch
1154	4	13.8	199	1	NHAA_RHOSO	Q53118	rhodococcus
1155	4	13.8	199	1	NP25_MOUSE	O9rlq8	mus musculus
1156	4	13.8	199	1	SODC_HAEDU	O59452	haemophilus
1157	4	13.8	199	1	TRPF_LACCA	P17218	lactobacilli
1158	4	13.8	199	1	VNSC_P13H4	P06165	human parai
1159	4	13.8	199	1	Y171_UREPA	Q9pqq2	ureaplasma
1160	4	13.8	200	1	ATP4_IPOBA	Q40089	ipomoea bat
1161	4	13.8	200	1	RUVA_CHLTR	O84509	chlamydia t
1162	4	13.8	200	1	Y873_TREPA	O83843	treponema p
1163	4	13.8	200	1	YCF4_LORJA	Q9bbr9	lotus japon
1164	4	13.8	200	1	YE01_MYCTU	P71669	mycobacteri
1165	4	13.8	200	1	YOR1_COYMY	P19200	commelina y
1166	4	13.8	201	1	ABP1_MAIZE	P13689	zea mays (m
1167	4	13.8	201	1	ARAS_ARAHY	P04149	arachis hyp
1168	4	13.8	201	1	HS73_PLAFA	P12078	plasmodium
1169	4	13.8	201	1	LEUD_ECOLI	P30126	escherichia
1170	4	13.8	201	1	LEUD_SALTY	P04787	salmonella
1171	4	13.8	201	1	MOBA_MYCTU	O53180	mycobacteri
1172	4	13.8	201	1	RL15_QUESU	O82712	quercus sub
1173	4	13.8	201	1	VID3_AGRV6	P09816	agrobacteri
1174	4	13.8	201	1	YAWC_SCHPO	Q10186	schizosacch
1175	4	13.8	201	1	YAW3_YEAST	Q44772	saccharomyc
1176	4	13.8	202	1	COAE_VIBCH	Q9kpe3	vibrio chol
1177	4	13.8	202	1	HOXM_ALCEU	P31909	alcaligenes
1178	4	13.8	202	1	RH01_SCHPO	Q09914	schizosacch
1179	4	13.8	203	1	AR61_HUMAN	Q15041	homo sapien
1180	4	13.8	203	1	AR61_MOUSE	Q91kw0	mus musculus
1181	4	13.8	203	1	RL15_ASFNG	O13418	aspergillus
1182	4	13.8	203	1	RPOC_PROSC	P42078	prochloron
1183	4	13.8	203	1	UL71_HSV7J	P52474	human herpe
1184	4	13.8	204	1	RIB3_SCHPO	O60181	schizosacch
1185	4	13.8	204	1	TRPE_SULSO	P50386	sulfolobus
1186	4	13.8	204	1	Y370_HAEIN	P43989	haemophilus
1187	4	13.8	204	1	Y4DW_RHLSN	P55422	rhizobium s
1188	4	13.8	204	1	YPOB_STYLE	P14289	stylonychia
1189	4	13.8	205	1	CEA8_ECOLI	P09882	escherichia
1190	4	13.8	205	1	DUT_RAT	P70583	rattus norv
1191	4	13.8	205	1	LEUD_BUCTS	Q31294	buchnera ap
1192	4	13.8	205	1	LUXR_VIBHA	P21308	vibrio harv
1193	4	13.8	206	1	COAE_PASNO	Q9cpf5	pasteurella
1194	4	13.8	206	1	HIT_HUMAN	P22492	homo sapien
1195	4	13.8	206	1	RRMJ_BUCAI	P57463	buchnera ap
1196	4	13.8	206	1	SC66_YEAST	P33754	saccharomyc
1197	4	13.8	206	1	Y140_METJA	O57605	methanococ
1198	4	13.8	206	1	YC10_ECOLI	P45847	escherichia
1199	4	13.8	206	1	YIG2_HAEIN	P44842	haemophilus
1200	4	13.8	207	1	CL13_HUMAN	O95833	homo sapien

ALIGNMENTS

RESULT 1

SELP_HUMAN

ID SELP_HUMAN STANDARD; PRT; 381 AA.

AC P49908;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Selenoprotein P precursor (Sep).

GN SEPP1 OR SELP.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

[1]

SEQUENCE FROM N.A.

RP TISSUE=Liver, and Heart;

RC MEDLINE=93133823; PubMed=8421687;

RA Hill K.E., Lloyd R.S., Burk R.F.;

RT "Conserved nucleotide sequences in the open reading frame and 3'

RT untranslated region of selenoprotein P mRNA.";

proc. Natl. Acad. Sci. U.S.A. 90:537-541(1993).

[2]

PARTIAL SEQUENCE.

RP MEDLINE=94191007; PubMed=8142465;

RA Aakesson B., Bellaw T., Burk R.F.;

RT "Purification of selenoprotein P from human plasma.";

RL Biochim. Biophys. Acta 1204:243-249(1994).

[3]

CHARACTERIZATION.

RP MEDLINE=98413836; PubMed=9735174;

RA Mostert V., Iombeck I., Abel J.;

RT "A novel method for the purification of selenoprotein P from human

RL plasma.";

Arch. Biochem. Biophys. 357:326-330(1998).

[4]

CHARACTERIZATION.

RP MEDLINE=20239644; PubMed=10775431;

RA Mostert V.;

RT "Selenoprotein P: properties, functions, and regulation.";

Arch. Biochem. Biophys. 376:433-438(2000).

[5]

REVIEW.

RP MEDLINE=95017128; PubMed=7931697;

RA Burk R.F., Hill K.E.;

RT "Selenoprotein P. A selenium-rich extracellular glycoprotein.";

J. Nutr. 124:1891-1897(1994).

-1- FUNCTION: MIGHT BE RESPONSIBLE FOR SOME OF THE EXTRACELLULAR

ANTIOXIDANT DEFENSE PROPERTIES OF SELENIUM OR MIGHT BE INVOLVED IN

THE TRANSPORT OF SELENIUM.

-1- SUBCELLULAR LOCATION: Extracellular.

-1- TISSUE SPECIFICITY: MADE IN THE LIVER AND HEART AND SECRETED INTO

THE PLASMA. IT IS ALSO FOUND IN THE KIDNEY.

-1- MISCELLANEOUS: THE SELENOCYSTEINES ARE ALL ENCODED BY THE OPAL

CODON, UGA.

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

or send an email to license@sib-sib.ch).

EMBL: Z11793; CAA77836.1; -

MIM: 601484; -

DR Glycoprotein; Signal; Selenium; Selenocysteine; Plasma.

KW SIGNAL 1 19

FT CHAIN 20 381 SELENOPROTEIN P.

FT DOMAIN 244 249 POLY-HIS.

```
FT SE_CYS 59 59
FT SE_CYS 300 300
FT SE_CYS 318 318
FT SE_CYS 330 330
FT SE_CYS 345 345
FT SE_CYS 352 352
FT SE_CYS 367 367
FT SE_CYS 376 376
FT SE_CYS 378 378
FT CARBOHYD 46 46
FT CARBOHYD 83 83
FT CARBOHYD 119 119
FT CARBOHYD 128 128
FT CARBOHYD 174 174
FT CARBOHYD 338 338
SQ SEQUENCE 381 AA; 42705 MW; 1A82FFA0BCD13CD7 CRC64;

Query Match 72.4%; Score 21; DB 1; Length 381;
Best Local Similarity 100.0%; Pred. No. 3.1e-15;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRCINQLLCKLPTDSELAAPS 21
Db 279 KRCINQLLCKLPTDSELAAPS 299

RESULT 2
SELP_MOUSE
ID SELP_MOUSE STANDARD; PRT; 380 AA.
AC P70274;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Selenoprotein P precursor (Sep).
GN SEPP1 OR SELP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SWISS WEBSTER / NIH;
RX MEDLINE=97434516; PubMed=9288402;
RA Steinert P., Ahrens M., Gross G., Flohe L.;
RT "cDNA and deduced polypeptide sequence of a mouse selenoprotein P.";
RL BioFactors 6:311-319(1997).
CC -1- FUNCTION: MIGHT BE RESPONSIBLE FOR SOME OF THE EXTRACELLULAR
CC ANTIOXIDANT DEFENSE PROPERTIES OF SELENIUM (BY SIMILARITY).
CC THE TRANSPORT OF SELENIUM (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: MADE IN THE LIVER AND SECRETED INTO THE
CC PLASMA (BY SIMILARITY).
CC -1- MISCELLANEOUS: THE SELENOCYSTEINES ARE ALL ENCODED BY THE OPAL
CC CODON, UGA.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X99807; CAA68140.1; -.
CC MGD: MGI:894288; Seppl.
CC Glycoprotein; Signal; Selenium; Selenocysteine; Plasma.
CC SIGNAL 1 19 BY SIMILARITY.
CC CHAIN 20 380 SELENOPROTEIN P.
CC SE_CYS 59 59
CC SE_CYS 259 259
CC SE_CYS 277 277
```

```
FT SE_CYS 318 318
FT SE_CYS 330 330
FT SE_CYS 352 352
FT SE_CYS 366 366
FT SE_CYS 368 368
FT SE_CYS 375 375
FT SE_CYS 377 377
FT CARBOHYD 83 83
FT CARBOHYD 176 176
FT CARBOHYD 195 195
FT CARBOHYD 365 365
FT CARBOHYD 370 370
SQ SEQUENCE 380 AA; 42236 MW; 72F7031941F47212 CRC64;

Query Match 31.0%; Score 9; DB 1; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CINQLLCKL 11
Db 281 CINQLLCKL 289

RESULT 3
SELP_RAT
ID SELP_RAT STANDARD; PRT; 385 AA.
AC P25236;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Selenoprotein P precursor (Sep).
GN SEPP1 OR SELP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-41; 267-287 AND 316-327.
RC TISSUE=Liver;
RX MEDLINE=91244760; PubMed=2037562;
RA Hill K.E., Lloyd R.S., Yang J.-G., Read R., Burk R.F.;
RT "The cDNA for rat selenoprotein P contains 10 TGA codons in the open
RT reading frame.";
RL J. Biol. Chem. 266:10050-10053(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Kidney;
RX MEDLINE=95364621; PubMed=7637580;
RA Saijoh K., Saito N., Lee M.J., Fujii M., Kobayashi T., Sumino K.;
RT "Molecular cloning of cDNA encoding a bovine selenoprotein P-like
RT protein containing 12 selenocysteines and a (His-Pro) rich domain
RT insertion, and its regional expression.";
RL Brain Res. Mol. Brain Res. 30:301-311(1995).
RN [3]
RP REVIEW.
RX MEDLINE=95017128; PubMed=7931697;
RA Burk R.F., Hill K.E.;
RT "Selenoprotein P, A selenium-rich extracellular glycoprotein.";
RL J. Nutr. 124:1891-1897(1994).
CC -1- FUNCTION: MIGHT BE RESPONSIBLE FOR SOME OF THE EXTRACELLULAR
CC ANTIOXIDANT DEFENSE PROPERTIES OF SELENIUM OR MIGHT BE INVOLVED IN
CC THE TRANSPORT OF SELENIUM.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: MADE IN THE LIVER AND SECRETED INTO THE
CC PLASMA.
CC -1- MISCELLANEOUS: THE SELENOCYSTEINES ARE ALL ENCODED BY THE OPAL
CC CODON, UGA.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```



```

RL Submitted (JAN-1989) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / FY1679;
RX MEDLINE=97051597; PubMed=8896274;
RA Saiz J.E., Buitrago M.J., Garcia R., Revuelta J.L., del Rey F.;
RT "The sequence of a 20.3 kb DNA fragment from the left arm of
RT Saccharomyces cerevisiae chromosome IV contains the KIN28, MSS2,
RT PHO2, POI3 and DUN1 genes, and six new open reading frames.";
RL Yeast 12:1077-1084(1996).
RN [4]
RN STRUCTURE, EXPRESSION, AND FUNCTION.
RP MEDLINE=89006272; PubMed=3049251;
RA Berben G., Legrain M., Hilger F.;
RT "Studies on the structure, expression and function of the yeast
RT regulatory gene PHO2.";
RL Gene 66:307-312(1988).
RN [5]
RN HOMEBOX DOMAIN.
RX MEDLINE=88210456; PubMed=2896548;
RA Buerklin T.R.;
RT "The yeast regulatory gene PHO2 encodes a homeo box.";
RL Cell 53:339-340(1988).
CC -1- FUNCTION: REGULATOR IN PHOSPHATE METABOLISM AND ACTS AS A
CC DEREPRESSOR OF ANOTHER CENTRAL REGULATOR PHO5. BINDS TO THE
CC UPSTREAM ACTIVATOR SEQUENCE (UAS) OF PHO5. IT ALSO BINDS TO THE
CC TRP4, HIS4, AND CYC1 PROMOTERS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: TWO REGIONS OF STRONG HOMOLOGY TO PHO4 ARE FOUND.
CC ALSO SIMILAR TO N.CRASSA NUCL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X05062; CAA28729.1; -
DR EMBL; M22259; AAA34866.1; -
DR EMBL; X95644; CAA64906.1; -
DR EMBL; Z74154; CAA98673.1; -
DR PIR; A25872; A25872.
DR HSP; P02836; 1HDD.
DR TRANSFAC; T00689; -
DR SGD; S0002264; GRF10.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
DR Homeobox; DNA-binding; Transcription regulation; Nuclear protein.
FT DOMAIN 23 52 GLN-RICH (INVOLVED IN TRANSCRIPTIONAL
FT ACTIVATION) (POTENTIAL).
FT DNA_BIND 77 136 HOMEBOX.
FT SEQUENCE 559 AA; 63390 MW; BE20E396D6DA0281 CRC64;
SQ
Query Match 24.1%; Score 7; DB 1; Length 559;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 LPTDSEL 17
| | | | |
Db 533 LPTDSEL 539
RESULT 6
RBS1_ARATH
ID RBS1_ARATH STANDARD; PRT; 180 AA.
AC P10795;
DT 01-JUL-1989 (Rel. 11, Created)

```

```

DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ribulose biphosphate carboxylase small chain 1A, chloroplast
DE precursor (EC 4.1.1.39) (RuBisCO small subunit 1A).
GN RBCS-1A OR ATSI1A.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=CV. COLOMBIA K85;
RA Kreebbers E., Seurinck J., Herdies L., Cashmore A.R., Timko M.P.;
RT "Four genes in two diverged subfamilies encode the ribulose-1,5-
RT biphosphate carboxylase small subunit polypeptides of Arabidopsis
RT thaliana.";
RL Plant Mol. Biol. 11:745-759(1988).
RN [2]
RN SEQUENCE OF 1-79 FROM N.A.
RP MEDLINE=92385772; PubMed=1515613;
RX Wong E.Y., Hironaka C.M., Fishhoff D.A.;
RT "Arabidopsis thaliana small subunit leader and transit peptide
RT enhance the expression of Bacillus thuringiensis proteins in
RT transgenic plants.";
RL Plant Mol. Biol. 20:81-93(1992).
CC -1- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF
CC D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
CC CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
CC THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
CC ACTIVE SITE.
CC -1- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + CO(2) = 2 3-
CC phospho-D-glycerate.
CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) =
CC 3-PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.
CC -1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- MISCELLANEOUS: THERE ARE FOUR GENES CODING FOR RBS IN ARABIDOPSIS
CC THALIANA.
CC -1- SIMILARITY: BELONGS TO THE RUBISCO SMALL CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X13611; CAA31948.1; ALT_SEQ.
DR EMBL; X68342; CAA48415.1; ALT_TERM.
DR PIR; S03720; RKWU1.
DR HSP; P00866; 4RUB.
DR Mendel; 269; ARATH; rbcS; 4.
DR InterPro; IPR000894; RuBisCO_small.
DR Pfam; PF00101; RuBisCO_small; 1.
DR PRINTS; PR00152; RUBISCO_SMALL.
DR ProDom; PD000290; RuBisCO_small; 1.
DR Photosynthesis; Carbon dioxide fixation; Photorespiration;
KW Lyase; oxidoreductase; Monooxygenase; Chloroplast; Transit peptide;
KW Multigene family.
FT TRANSIT 1 55 CHLOROPLAST.
FT CHAIN 56 180 RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL
FT CHAIN 1A.
FT SEQUENCE 180 AA; 20202 MW; 5DFC6B76D61F8F7 CRC64;
SQ
Query Match 20.7%; Score 6; DB 1; Length 180;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13 TDSELA 18

```



```
Db 77 TDSELA 82
|||||
RESULT 7
KTHY_PYRHO
ID KTHY_PYRHO STANDARD: PRT; 205 AA.
AC O59366;
DT 15-JUL-1999 (Rel. 38, Last Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable thymidylate kinase (EC 2.7.4.9) (dTMP kinase).
GN TMK OR PH1695.
OS Pyrococcus horikoshii.
OC Archaea: Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=OT3;
RX MEDLINE=90344137; PubMed=9679194;
RA Kawarabayashi Y., Savada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yanamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Rudooh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
RL DNA Res. 5:55-76(1998).
CC -!- CATALYTIC ACTIVITY: ATP + thymidine 5'-phosphate = ADP + thymidine
CC 5'-diphosphate.
CC -!- PATHWAY: SYNTHESIS OF dTTP FROM dTMP.
CC -!- SIMILARITY: BELONGS TO THE THYMIDYLATE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP000007; BAA30808.1; -.
DR HSSP; P37345; STMP.
DR InterPro; IPR000062; Thymidylate_kin.
DR Pfam; PF02223; Thymidylate_kin; 1.
DR PROSITE; PS01331; THYMIDYLATE_KINASE; 1.
KW Transferase; Kinase; Nucleotide biosynthesis; ATP-binding;
KW Complete proteome.
FT NP_BIND 10 17 ATP (POTENTIAL).
SQ SEQUENCE 205 AA; 23399 MW; E7F623520E97C11C CRC64;

Query Match 20.7%; Score 6; DB 1; Length 205;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSEL 17
|||||
DB 40 PTDSEL 45

RESULT 8
YD93_RHILO
ID YD93_RHILO STANDARD: PRT; 302 AA.
AC Q98KN6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein M111393.
GN M111393.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MAFF303099; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
CC -!- SIMILARITY: BELONGS TO THE PHZF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP002997; BAB48778.1; -.
DR InterPro; IPR003719; PhzC_PhzF.
DR Pfam; PF02567; PhzC-PhzF; 1.
KW Hypothetical protein; Oxidoreductase; Complete proteome.
SQ SEQUENCE 302 AA; 32644 MW; DB460394C4BA8298 CRC64;

Query Match 20.7%; Score 6; DB 1; Length 302;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SELAPR 20
|||||
DB 189 SELAPR 194

RESULT 9
YAHB_ECOLI
ID YAHB_ECOLI STANDARD: PRT; 310 AA.
AC P77700;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical transcriptional regulator yahb.
GN YAHB OR B0316.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,
RA Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,
RA Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AF000138; AAC73419.1; -;
 DR EMBL; U73857; AAB18042.1; -;
 DR Ecogene; EG13586; yahB.
 DR InterPro; IPR002197; HTH_Fis.
 DR InterPro; IPR000847; HTH_LysR.
 DR Pfam; PF00126; HTH_1.1;
 DR PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
 KW Hypothetical protein; Transcription regulation; DNA-binding;
 FT Complete proteome.
 KW DNA_BIND 22 42 H-T-H MOTIF (POTENTIAL).
 SQ SEQUENCE 310 AA; 34866 MW; A3EB895E7F69D60C CRC64;

Query Match 20.7%; Score 6; DB 1; Length 310;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 CINQLL 8
 DB 100 CINQLL 105

RESULT 10
 TSPB_CAEEL STANDARD; PRT; 359 AA.
 AC Q20351; 2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Putative protein-tyrosine sulfotransferase (EC 2.8.2.20)
 DE (Tyrosylprotein sulfotransferase) (TPST).
 GN F4269.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Taich A.;
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CATALYZES THE O-SULFATION OF TYROSINE RESIDUES WITHIN
 CC ACIDIC MOTIFS OF POLYPEPTIDES (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: 3'-phosphoadenylylsulfate + protein tyrosine -
 CC adenosine 3',5'-bisphosphate + protein tyrosine-O-sulfate.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Golgi membrane (By
 CC similarity).
 CC -1- SIMILARITY: BELONGS TO THE PROTEIN SULFOTRANSFERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U00051; AAA91354.1; -;
 DR WormPep; F4269.8; CE07235.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransfer; 1.
 KW Hypothetical protein; Transferase; Transmembrane; Glycoprotein;
 KW Signal-anchor.
 FT DOMAIN 1 27 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 28 48 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 FT (POTENTIAL).

FT DOMAIN 49 359 LUMENAL, CATALYTIC (POTENTIAL).
 SQ SEQUENCE 359 AA; 41469 MW; 9FC2F44539B8CC53 CRC64;

Query Match 20.7%; Score 6; DB 1; Length 359;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 SELAPR 20
 DB 161 SELAPR 166

RESULT 11
 NCF1_MOUSE STANDARD; PRT; 390 AA.
 ID Q09014; 070144; O9J134;
 AC Q09014; 070144; O9J134;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Neutrophil cytosol factor 1 (NCF-1) (Neutrophil NADPH oxidase factor
 DE 1) (47 kDa neutrophil oxidase factor) (p47-phox) (NCF-47K).
 GN NCF1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Macrophage;
 RX MEDLINE=94164697; PubMed=8119734;
 RA Jackson S.H., Malech H.L., Kozak C.A., Lomax K.J., Gallin J.I.,
 RA Holland S.M.;
 RT "Cloning and functional expression of the mouse homologue of
 RT p47phox";
 RL Immunogenetics 39:272-275(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leukemia;
 RX MEDLINE=98149672; PubMed=9490028;
 RA Mizuki K., Kadamatsu K., Hata K., Ito T., Fan O.-W., Kage Y.,
 RA Fukumaki Y., Sakaki Y., Takeshige K., Sumimoto H.;
 RT "Functional modules and expression of mouse p40(phox) and p67(phox),
 RT SH3-domain-containing proteins involved in the phagocyte NADPH oxidase
 RT complex";
 RL Eur. J. Biochem. 251:573-582(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Green E.D.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: NCF2, NCF1, AND A MEMBRANE BOUND CYTOCHROME B558 ARE
 CC REQUIRED FOR ACTIVATION OF THE LATENT NADPH OXIDASE (NECESSARY
 CC FOR SUPEROXIDE PRODUCTION).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: CONTAINS 2 SH3 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 PX DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L11455; AAA50469.1; -;
 DR EMBL; AB002663; BAA25649.1; -;
 DR EMBL; AF267747; AAF90134.1; -;
 DR MGD; MGI:97283; Ncf1.
 DR InterPro; IPR001655; p47PHOX.
 DR InterPro; IPR001683; PX.
 DR InterPro; IPR001452; SH3.
 DR Pfam; PF00787; PX; 1.

```
DR Pfam; PF00018; SH3; 2.
DR PRINTS; PR00498; P47PHOX.
DR PRINTS; PR00452; SH3DOMAIN.
DR SMART; SM00312; PX; 1.
DR SMART; SM00326; SH3; 2.
DR PROSITE; PSS0002; SH3; 2.
KW SH3 domain; Repeat.
FT DOMAIN 9 128 PX.
FT DOMAIN 211 254 ASP/GLU-RICH (HIGHLY ACIDIC).
FT DOMAIN 292 390 ARG/LYS-RICH (HIGHLY BASIC).
FT DOMAIN 162 215 SH3 1.
FT DOMAIN 226 285 SH3 2.
FT CONFLICT 161 161 Y -> H (IN REF. 1).
FT CONFLICT 343 345 GOL -> RAA (IN REF. 1).
FT CONFLICT 344 344 Q -> P (IN REF. 3).
FT CONFLICT 344 344 Q -> P (IN REF. 3).
SQ SEQUENCE 390 AA; 44698 MW; A383DB953839CFB CRC64;

Query Match 20.7%; Score 6; DB 1; Length 390;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KLPTDS 15
Db 126 KLPTDS 131

RESULT 12
ID NCF1_BOVIN STANDARD; PRT; 392 AA.
AC Q07774;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Neutrophil cytosol factor 1 (NCF-1) (Neutrophil NADPH oxidase factor
DE 1) (47 kDa neutrophil oxidase factor) (p47-phox) (NCF-47k).
GN NCF1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Bunker P L., Siemsen D., Quinn M.T.;
RT "Cloning and expression of bovine NADPH oxidase proteins p47-phox and
RT p67-phox."
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NCF2, NCF1, AND A MEMBRANE BOUND CYTOCHROME B558 ARE
CC REQUIRED FOR ACTIVATION OF THE LATENT NADPH OXIDASE (NECESSARY
CC FOR SUPEROXIDE PRODUCTION).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: CONTAINS 2 SH3 DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 PX DOMAIN.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AF079302; AAC82462.1; -
InterPro; IPR001655; P47PHOX.
InterPro; IPR001683; PX.
InterPro; IPR001452; SH3.
Pfam; PF00787; PX; 1.
Pfam; PF00018; SH3; 2.
PRINTS; PR00498; P47PHOX.
PRINTS; PR00452; SH3DOMAIN.
SMART; SM00312; PX; 1.
SMART; SM00326; SH3; 2.
```

```
DR PROSITE; PSS0002; SH3; 2.
KW SH3 domain; Repeat.
FT DOMAIN 9 128 PX.
FT DOMAIN 211 254 ASP/GLU-RICH (HIGHLY ACIDIC).
FT DOMAIN 292 392 ARG/LYS-RICH (HIGHLY BASIC).
FT DOMAIN 156 215 SH3 1.
FT DOMAIN 226 285 SH3 2.
SQ SEQUENCE 392 AA; 45345 MW; 79199BD86AE80DB7 CRC64;

Query Match 20.7%; Score 6; DB 1; Length 392;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KLPTDS 15
Db 126 KLPTDS 131

RESULT 13
ID MPK1_XENLA STANDARD; PRT; 394 AA.
AC Q05116;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Dual specificity mitogen-activated protein kinase kinase 1
DE (EC 2.7.1.-) (MAP kinase kinase 1) (MAPKK 1) (ERK activator kinase 1)
DE (MAPK/ERK kinase 1) (MEK1).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE-93178455; PubMed-8440264;
RA Kosako H., Nishida E., Gotoh Y.;
RT "cDNA cloning of MAP kinase kinase reveals kinase cascade pathways in
RT yeasts to vertebrates."
RL EMBO J. 12:787-794(1993).
CC -!- FUNCTION: CATALYZES THE CONCOMITANT PHOSPHORYLATION OF A THREONINE
CC AND A TYROSINE RESIDUE IN A THR-GLU-TYR SEQUENCE LOCATED IN MAP
CC KINASES.
CC -!- TISSUE SPECIFICITY: IS EXPRESSED IN THE CENTRAL NERVOUS SYSTEM,
CC KIDNEYS, LIVER, INTESTINE AND THE HEMATOPOIETIC SYSTEM.
CC -!- PM: MAPKK IS ITSELF DEPENDENT ON SER/THR PHOSPHORYLATION FOR
CC ACTIVITY CATALYZED BY MAP KINASE KINASE KINASES (KAP OR MOS).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE SUBFAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; D13700; BAA02860.1; -
PIR; S36186; S36186.
HSSP; P11362; IFGK.
InterPro; IPR000719; Euk_pkinase.
InterPro; IPR002290; Ser_thr_kinase.
Pfam; PF00069; pkinase; 1.
SMART; SM00220; S_TKc; 1.
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
KW ATP-binding; Phosphorylation.
FT INIT_MET 0
FT DOMAIN 67 362 PROTEIN KINASE.
```

```

FT NP_BIND 73 81 ATP (BY SIMILARITY).
FT BINDING 96 96 BY SIMILARITY).
FT ACT_SITE 189 189 BY SIMILARITY.
FT MOD_RES 217 217 PHOSPHORYLATION (BY RAF) (BY SIMILARITY).
FT MOD_RES 221 221 PHOSPHORYLATION (BY RAF) (BY SIMILARITY).
FT DOMAIN 261 308 PRO-RICH.
SQ SEQUENCE 394 AA; 43611 MW; 102B9BF954BA2053 CRC64;

Query Match 20.7%; Score 6; DB 1; Length 394;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SELAPR 20
Db 285 SELAPR 290
|||||

RESULT 14
LACY_KLEPN
ID LACY_KLEPN STANDARD; PRT; 416 AA.
AC P18817;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Lactose permease (Lactose-proton symport).
GN LACY.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M5A1;
RX MEDLINE=89050993; PubMed=30565524;
RA McMorow I., Chin D.T., Fiebig K., Pierce J.L., Wilson D.M.,
RA Reeve E.C.R., Wilson T.H.;
RT "The lactose carrier of Klebsiella pneumoniae M5a1; the physiology of
RT transport and the nucleotide sequence of the lacY gene.";
RL Biochim. Biophys. Acta 945:315-323(1988).
RN [2]
RP SEQUENCE OF 1-131 FROM N.A.
RC STRAIN=TI7R1;
RX MEDLINE=85289025; PubMed=3897196;
RA Buvinger W.E., Riley M.;
RT "Nucleotide sequence of Klebsiella pneumoniae lac genes.";
RL J. Bacteriol. 163:850-857(1985).
CC -1- FUNCTION: RESPONSIBLE FOR TRANSPORT OF BETA-GALACTOSIDES INTO THE
CC CELL, WITH THE CONCOMITANT IMPORT OF A PROTON (SYMPORT SYSTEM).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -1- SIMILARITY: BELONGS TO THE LACY/RAFB FAMILY OF PERMEASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14154; CAA32366.1;
DR EMBL; M11441; AAA25083.1; ALT_INIT.
DR PIR; JT0487; JT0487.
DR PIR; C24925; C24925.
DR InterPro; IPR000576; Lacy_symp.
DR Pfam; PF01306; Lacy_symp.1.
DR PRINTS; PR00174; LACYSMPORT.
DR PROSITE; PS00896; LACY_1; 1.
DR PROSITE; PS00897; LACY_2; 1.
KW Transport; Sugar transport; Symport; Inner membrane; Transmembrane.
FT DOMAIN 1 13 CYTOPLASMIC.
FT TRANSMEM 14 34 1 (POTENTIAL).
FT DOMAIN 35 50 PERIPLASMIC (POTENTIAL).

```

```

FT TRANSMEM 51 71 2 (POTENTIAL).
FT DOMAIN 72 80 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 81 101 3 (POTENTIAL).
FT DOMAIN 102 107 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 108 130 4 (POTENTIAL).
FT DOMAIN 131 149 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 150 170 5 (POTENTIAL).
FT DOMAIN 171 172 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 173 193 6 (POTENTIAL).
FT DOMAIN 194 223 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 224 244 7 (POTENTIAL).
FT DOMAIN 245 267 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 268 288 8 (POTENTIAL).
FT DOMAIN 289 295 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 296 316 9 (POTENTIAL).
FT DOMAIN 317 319 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 320 339 10 (POTENTIAL).
FT DOMAIN 340 353 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 354 374 11 (POTENTIAL).
FT DOMAIN 375 384 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 385 405 12 (POTENTIAL).
FT DOMAIN 406 416 CYTOPLASMIC.
FT VARIANT 3 3 L -> F (IN STRAIN TI7R1).
FT VARIANT 15 15 I -> V (IN STRAIN TI7R1).
FT VARIANT 18 18 M -> L (IN STRAIN TI7R1).
FT VARIANT 25 25 Y -> H (IN STRAIN TI7R1).
FT VARIANT 41 41 E -> D (IN STRAIN TI7R1).
FT VARIANT 56 56 C -> S (IN STRAIN TI7R1).
FT VARIANT 71 71 I -> M (IN STRAIN TI7R1).
FT VARIANT 87 87 I -> V (IN STRAIN TI7R1).
FT VARIANT 109 109 M -> I (IN STRAIN TI7R1).
FT VARIANT 112 112 A -> S (IN STRAIN TI7R1).
FT VARIANT 117 117 V -> I (IN STRAIN TI7R1).
FT VARIANT 125 131 SRSGAVE -> TAPGVGS (IN STRAIN TI7R1).
SQ SEQUENCE 416 AA; 46220 MW; A37D8BA4C38C467B CRC64;

Query Match 20.7%; Score 6; DB 1; Length 416;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SELAPR 20
Db 4 SELAPR 9
|||||

RESULT 15
AMYA_ASAPW
ID AMYA_ASAPW STANDARD; PRT; 498 AA.
AC Q02905;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Alpha-amylase A precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase A).
GN AMYA.
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=105351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UVK143F;
RX MEDLINE=90254827; PubMed=2340591;
RA Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,
RA Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,
RA Berka R.M.;
RT "Cloning, characterization, and expression of two alpha-amylase genes
RT from Aspergillus niger var. awamori.";
RL Curr. Genet. 17:203-212(1990).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.

```



```

DR PDB; 2TAA; 15-OCT-89.
DR PDB; 6TAA; 31-OCT-93.
DR PDB; 7TAA; 25-NOV-98.
DR GlycoSuiteDB; P10529;
DR InterPro; IPR000461; Alpha_amylase.
DR Pfam; PF00128; alpha-amylase; 1.
KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
KW Glycoprotein; Signal; Multigene family; 3D-structure.
FT SIGNAL 1 21
FT CHAIN 22 499 ALPHA-AMYLASE A.
FT ACT_SITE 227 227
FT ACT_SITE 231 231
FT ACT_SITE 251 251
FT ACT_SITE 318 318
FT DISULFID 51 59
FT DISULFID 171 185
FT DISULFID 261 304
FT DISULFID 461 496
FT CARBOHYD 218 218
FT VARIANT 56 56
FT VARIANT 172 172
FT CONFLICT 93 94
FT CONFLICT 106 106
FT CONFLICT 184 184
FT CONFLICT 195 195
FT CONFLICT 255 255
FT CONFLICT 291 291
FT CONFLICT 345 345
FT CONFLICT 370 370
FT CONFLICT 406 406
FT CONFLICT 448 448
FT CONFLICT 497 497
FT HELIX 24 27
FT TURN 28 29
FT STRAND 32 35
FT HELIX 37 40
FT STRAND 42 42
FT TURN 43 44
FT HELIX 53 55
FT STRAND 61 61
FT HELIX 63 67
FT TURN 68 69
FT HELIX 70 74
FT TURN 75 77
FT STRAND 80 83
FT STRAND 87 89
FT STRAND 94 94
FT TURN 95 96
FT STRAND 97 97
FT TURN 100 101
FT STRAND 105 111
FT TURN 113 115
FT TURN 118 119
FT HELIX 120 130
FT TURN 131 132
FT STRAND 134 139
FT STRAND 143 143
FT STRAND 146 146
FT HELIX 150 152
FT TURN 155 156
FT STRAND 159 159
FT HELIX 164 166
FT STRAND 167 167
FT STRAND 172 172
FT HELIX 179 184
FT STRAND 186 188
FT STRAND 193 194
FT STRAND 196 197
FT TURN 199 200
FT HELIX 202 219
FT TURN 220 220
FT STRAND 223 226

N-LINKED (GLCNAC. .).
/FTIG-CAR_000125.
Q -> R (IN AMY3).
F -> L (IN AMY3).
TT -> DC (IN REF. 5).
Q -> T (IN REF. 5).
D -> Y (IN REF. 3).
P -> L (IN REF. 3).
G -> V (IN REF. 3).
D -> H (IN REF. 4).
I -> L (IN REF. 5).
L -> A (IN REF. 4).
WPIY -> PYI (IN REF. 5).
G -> S (IN REF. 5).
S -> SD (IN REF. 5 AND 7).

Query Match 20.7%; Score 6; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSEL 17
Db 374 PTDSEL 379

RESULT 17
AMYB_ASPAW
ID AMYB_ASPAW STANDARD; PRT; 499 AA.
AC Q02906;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Alpha-amylase B precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase B).
GN AMYB.
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Euryotales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=105351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UVK143F;
RX MEDLINE=90254827; PubMed=2340591;
RA Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,
RA Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,
RA Berka R.M.;
RT "Cloning, characterization, and expression of two alpha-amylase genes
from Aspergillus niger var. awamori.";
RL Curr. Genet. 17:203-212(1990).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -!- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X52756; CAA36967.1; -.
CC HSSP; P10529; 7TAA.

```

DR InterPro; IPR000461; Alpha-amylase.
 DR Pfam; PF00128; alpha-amylase; 1.
 KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
 KW Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 499 ALPHA-AMYLASE B.
 FT ACT_SITE 227 227 BY SIMILARITY.
 FT ACT_SITE 231 231 BY SIMILARITY.
 FT ACT_SITE 251 251 BY SIMILARITY.
 FT ACT_SITE 318 318 BY SIMILARITY.
 FT DISULFID 51 59 BY SIMILARITY.
 FT DISULFID 171 185 BY SIMILARITY.
 FT DISULFID 261 304 BY SIMILARITY.
 FT DISULFID 461 496 BY SIMILARITY.
 FT CARBOHYD 218 218 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 499 AA; 54921 MW; 740B96B11BC01A8A CRC64;

Query Match 20.7%; Score 6; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSEL 17
 DB 374 PTDSEL 379

RESULT 18
 AMT_ASPSH
 ID AMY_ASPSH STANDARD; PRT; 499 AA.
 AC P30292;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
 DE glucanohydrolase).
 GN AMY.
 OS Aspergillus shirousami.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5070;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92323146; PubMed-1368777;
 RA Shibuya I., Tamura G., Ishikawa T., Hara S.;
 RT "Cloning of the alpha-amylase cDNA of Aspergillus shirousami and its
 RT expression in Saccharomyces cerevisiae.";
 RL Biosci. Biotechnol. Biochem. 56:174-179(1992).
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides.
 CC -!- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D10461; BAA01255.1; -.
 CC PIR; JS0663; JS0663.
 CC HSP; P10529; 7TAA.
 DR InterPro; IPR000461; Alpha-amylase.
 DR Pfam; PF00128; alpha-amylase; 1.
 KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
 KW Glycoprotein; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 499 ALPHA-AMYLASE.
 FT ACT_SITE 227 227
 FT ACT_SITE 231 231 BY SIMILARITY.

FT ACT_SITE 251 251 BY SIMILARITY.
 FT ACT_SITE 318 318 BY SIMILARITY.
 FT DISULFID 51 59 BY SIMILARITY.
 FT DISULFID 171 185 BY SIMILARITY.
 FT DISULFID 261 304 BY SIMILARITY.
 FT DISULFID 461 496 BY SIMILARITY.
 FT CARBOHYD 218 218 N-LINKED (GLCNAC...) (BY SIMILARITY).
 SQ SEQUENCE 499 AA; 54852 MW; 1FB7AE50DA01C03F CRC64;

Query Match 20.7%; Score 6; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSEL 17
 DB 374 PTDSEL 379

RESULT 19
 DGTP_SALTY
 ID DGTP_SALTY STANDARD; PRT; 504 AA.
 AC P40733;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Deoxyguanosinetriphosphate triphosphohydrolase (EC 3.1.5.1) (dGTPase)
 DE (dGTP triphosphohydrolase).
 GN DGT OR STM0208.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE-21534948; PubMed-11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Willson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 RT LT2.";
 RL Nature 413:852-856(2001).
 RN [2]
 RP SEQUENCE OF 418-504 FROM N.A.
 RC STRAIN-C5;
 RX MEDLINE-91251770; PubMed-1645840;
 RA Johnson K., Charles I., Dougan G., Pickard D., O'Gaora P., Costa G.,
 RA Ali T., Miller I., Hormaeche C.;
 RT "The role of a stress-response protein in Salmonella typhimurium
 RT virulence.";
 RL Mol. Microbiol. 5:401-407(1991).
 CC -!- FUNCTION: DGTPASE PREFERENTIALLY HYDROLYZES DGTP OVER THE OTHER
 CC CANONICAL NTPS (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: dGTP + H(2)O = deoxyguanosine + triphosphate.
 CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE DGTPASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE008704; AAL19172.1; -.
 CC EMBL; X54548; -; NOT_ANNOTATED_CDS.
 CC StyGene; SGI0473; dgt.
 KW Hydrolase; Complete proteome.
 FT INIT_MET 0 BY SIMILARITY.

CC -!- TISSUE SPECIFICITY: EXPRESSED IN ROOTS AND ROSETTES.
CC -!- PTM: AUTOPHOSPHORYLATED ON SERINE AND THREONINE RESIDUES.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SER/THR
CC FAMILY OF PROTEIN KINASES.
CC -----
CC -!- SIMILARITY: CONTAINS 18 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M84660; AAA32859.1; -.
DR EMBL; AL021749; CAAL6889.1; -.
DR EMBL; AL161572; CABY9651.1; -.
DR HSSP; P00523; 2PTK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00560; LRR; 20.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00370; LRR; 17.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW Transmembrane; Serine/threonine-protein kinase; ATP-binding;
KW Repeat; Signal.
FT SIGNAL 1 14 POTENTIAL.
FT CHAIN 15 999 RECEPTOR-LIKE PROTEIN KINASE 5.
FT DOMAIN 15 621 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 622 999 POTENTIAL.
FT DOMAIN 642 999 CYTOPLASMIC (POTENTIAL).
FT REPEAT 88 112 LRR 1.
FT REPEAT 114 138 LRR 2.
FT REPEAT 139 161 LRR 3.
FT REPEAT 163 186 LRR 4.
FT REPEAT 188 211 LRR 5.
FT REPEAT 235 261 LRR 6.
FT REPEAT 263 283 LRR 7.
FT REPEAT 284 306 LRR 8.
FT REPEAT 307 330 LRR 9.
FT REPEAT 332 353 LRR 10.
FT REPEAT 354 378 LRR 11.
FT REPEAT 402 426 LRR 12.
FT REPEAT 427 450 LRR 13.
FT REPEAT 452 474 LRR 14.
FT REPEAT 498 522 LRR 15.
FT REPEAT 524 547 LRR 16.
FT REPEAT 549 567 LRR 17.
FT REPEAT 568 598 LRR 18.
FT DOMAIN 683 968 PROTEIN KINASE.
FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 452 452 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 576 576 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NP_BIND 689 697 ATP (BY SIMILARITY).
FT BINDING 711 711 ATP (BY SIMILARITY).
FT ACT_SITE 819 819 K-E: LOSS OF CATALYTIC ACTIVITY.
FT MUTAGEN 711 711 K-E: F5793D899EA0C6A7 CRC64;
SQ SEQUENCE 999 AA; 109095 MW; F5793D899EA0C6A7 CRC64;

Query Match

20.7%; Score 6; DB 1; Length 999;

Best Local Similarity 100.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 12 PTDSEL 17
Db 897 PTDSEL 902
RESULT 22
RPOP_AGABT
ID RPOP_AGABT STANDARD; PRT; 1102 AA.
AC P33539;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Probable DNA-directed RNA polymerase (EC 2.7.7.6).
OS Agaricus bisporus.
OC Mitochondrion.
OC Plasmid pEM.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Agaricaceae; Agaricus.
OX NCBI_taxid=5343;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 24666 / AG4;
RX MEDLINE=91347410; PubMed=1879001;
RA Robison M.M., Royer J.C., Horgen P.A.;
RT "Homology between mitochondrial DNA of Agaricus bisporus and an
RT internal portion of a linear mitochondrial plasmid of Agaricus
RT bisporus.";
RL Curr. Genet. 19:495-502(1991).
CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC [RNA](N).
CC -!- SIMILARITY: BELONGS TO THE PHAGE AND MITOCHONDRIAL RNA POLYMERASES
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X63075; CAA44799.1; -.
DR PIR; S28104; S28104.
DR InterPro; IPR002092; RNA_pol_phage.
DR Pfam; PF00940; RNA_pol; 1.
DR PROSITE; PS00489; RNA_POL_PHAGE_2; 1.
DR PROSITE; PS00900; RNA_POL_PHAGE_1; 1.
KW Transferase; Transcription; DNA-directed RNA polymerase;
KW Mitochondrion; Plasmid.
FT ACT_SITE 734 734 BY SIMILARITY.
FT ACT_SITE 804 804 BY SIMILARITY.
FT ACT_SITE 980 980 BY SIMILARITY.
SQ SEQUENCE 1102 AA; 126664 MW; F72FDD308D1A85B4 CRC64;
Query Match 20.7%; Score 6; DB 1; Length 1102;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 10 KLPTDS 15
Db 933 KLPTDS 938
RESULT 23
POLG_TMEVD
ID POLG_TMEVD STANDARD; PRT; 2301 AA.

AC P13899; Q88564; Q88565; Q88566; Q88567; Q88568; Q88569; Q88570;
AC Q88571; Q88572; Q88573; Q88574; Q88580;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C, P3A; Genome-linked protein VPg; Picornain 3C
DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D
DE (EC 2.7.7.48)].
OS Theiler's murine encephalomyelitis virus (strain DA) (TMEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Cardiovirus;
OX NCBI_TaxID=12126;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=88206072; PubMed=2834872;
RX Ohara Y., Stein S., Fu J., Stillman L., Klamon L., Roos R.P.;
RA "Molecular cloning and sequence determination of DA strain of
RT Theiler's murine encephalomyelitis viruses.";
RL Virology 164:245-255(1988).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX MEDLINE=92196057; PubMed=1549565;
RA Grant R.A., Filman D.J., Fujinami R.S., Icenogle J.P., Hogle J.M.;
RT "Three-dimensional structure of Theiler virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2061-2065(1992).
CC -1- FUNCTION: IT IS THOUGHT THAT THE P2C PROTEIN ATTACHES TO VESICULAR
CC MEMBRANES AND IS ASSOCIATED WITH VIRAL RNA SYNTHESIS.
CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: CLOSELY RELATED TO ENCEPHALOMYOCARDITIS VIRUS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M20301; AAA47928.1; -;
DR PIR; A31228; GNNTN.
DR PDB; 1TME; 31-JAN-94.
DR MEROPS; C03.009; -;
DR MEROPS; U29.001; -;
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR001676; Rhv.
DR Pfam; PF00073; rhv; 3.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
KW Polypeptide; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
KW 3D-structure.
FT PROPEP 1 76 LEADER PEPTIDE.
FT CHAIN 77 147 COAT PROTEIN VP4 (PIA).
FT CHAIN 148 414 COAT PROTEIN VP2 (PIB).
FT CHAIN 415 646 COAT PROTEIN VP3 (PIC).
FT CHAIN 647 920 COAT PROTEIN VP1 (PID).
FT CHAIN 921 1062 CORE PROTEIN P2A.
FT CHAIN 1063 1189 CORE PROTEIN P2B.
FT CHAIN 1190 1515 CORE PROTEIN P2C.
FT CHAIN 1516 1603 CORE PROTEIN P3A.
FT CHAIN 1604 1623 GENOME-LINKED PROTEIN VPg (P3B).
FT CHAIN 1624 1840 PICORNAIN 3C.
FT CHAIN 1841 2301 RNA-DIRECTED RNA POLYMERASE P3D.
FT LIPID 77 77 MYRISTATE (BY SIMILARITY).
FT ACT_SITE 1791 1791 PROTEASE (POTENTIAL).
FT ACT_SITE 1809 1809 PROTEASE (POTENTIAL).
FT TURN 96 97

FT HELIX 103 106
FT STRAND 109 110
FT STRAND 162 166
FT TURN 167 168
FT STRAND 169 173
FT STRAND 179 181
FT HELIX 182 184
FT TURN 193 194
FT STRAND 200 201
FT HELIX 204 206
FT STRAND 210 217
FT TURN 219 220
FT TURN 223 224
FT STRAND 226 231
FT TURN 232 232
FT HELIX 233 235
FT HELIX 238 247
FT TURN 248 249
FT STRAND 250 262
FT TURN 267 268
FT STRAND 270 279
FT TURN 280 280
FT STRAND 286 287
FT TURN 289 291
FT STRAND 294 295
FT TURN 297 298
FT STRAND 305 305
FT STRAND 324 324
FT TURN 325 326
FT HELIX 330 335
FT STRAND 338 342
FT TURN 343 345
FT STRAND 348 353
FT STRAND 362 362
FT STRAND 364 366
FT STRAND 367 367
FT STRAND 370 381
FT TURN 384 385
FT STRAND 390 406
FT TURN 423 426
FT TURN 430 431
FT STRAND 438 438
FT STRAND 444 445
FT TURN 449 450
FT STRAND 454 455
FT STRAND 457 457
FT HELIX 458 463
FT STRAND 466 467
FT TURN 470 471
FT TURN 473 474
FT STRAND 477 478
FT STRAND 480 484
FT STRAND 493 496
FT TURN 499 500
FT TURN 503 506
FT HELIX 508 513
FT TURN 514 515
FT STRAND 516 520
FT STRAND 523 529
FT TURN 533 534
FT STRAND 536 544
FT HELIX 554 557
FT TURN 558 559
FT STRAND 561 566
FT STRAND 572 577
FT STRAND 586 587
FT STRAND 601 611
FT TURN 614 615
FT STRAND 619 628
FT TURN 630 631
FT STRAND 633 637
FT STRAND 648 649
FT HELIX 651 653

```
FT TURN 661 664
FT TURN 676 677
FT STRAND 678 678
FT HELIX 679 683
FT STRAND 687 692
FT TURN 698 699
FT STRAND 706 707
FT TURN 708 709
FT STRAND 710 711
FT STRAND 713 715
FT STRAND 721 722
FT TURN 727 728
FT TURN 732 733
FT STRAND 737 738
FT STRAND 741 741
FT STRAND 744 744
FT STRAND 752 752
FT STRAND 755 755
FT STRAND 758 762
FT TURN 763 763
FT STRAND 766 779
FT STRAND 786 792
FT TURN 794 795
FT TURN 806 807
FT HELIX 810 812
FT STRAND 818 822
FT TURN 824 825
FT STRAND 828 833
FT STRAND 842 843
FT STRAND 847 847
FT STRAND 849 849
FT TURN 853 854
FT TURN 859 859
FT TURN 860 860
FT TURN 863 864
FT STRAND 869 874
FT STRAND 878 892
SQ SEQUENCE 2301 AA; 256159 MW; 0B6095DF153DBFDF CRC64;

Query Match 20.78; Score 6; DB 1; Length 2301;
Best Local Similarity 100.0%; Pred. No. 1.le+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 LCKLPT 13
Db 461 LCKLPT 466

RESULT 24
POLG_TWEVB STANDARD; PRT; 2303 AA.
AC P08544; Q88583; Q88584; Q88585; Q88586; Q88587; Q88588; Q88589;
AC Q88590; Q88591; Q88592;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; P3A; Genome-linked protein VPg; Picornain 3C
DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D
DE (EC 2.7.7.48)].
OS Theiler's murine encephalomyelitis virus (strain BeAn 8386) (TMEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Cardiovirus.
OX NCBI_TaxID=12125;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87198877; PubMed=3033278;
RA Pevear D.C., Calenoff M., Rozhon E., Lipton H.L.;
RT "Analysis of the complete nucleotide sequence of the picornavirus
RT Theiler's murine encephalomyelitis virus indicates that it is closely
RT related to cardioviruses.";
RL J. Virol. 61:1507-1516(1987).

[2]
RN X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RX MEDLINE=92196127; PubMed=1312722;
RA Luo M., He C., Toth K.S., Zhang C.X., Lipton H.L.;
RT "Three-dimensional structure of Theiler murine encephalomyelitis
RT virus (BeAn strain).";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2409-2413(1992).
CC -1- FUNCTION: IT IS THOUGHT THAT THE P2C PROTEIN ATTACHES TO VESICULAR
CC MEMBRANES AND IS ASSOCIATED WITH VIRAL RNA SYNTHESIS.
CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: CLOSELY RELATED TO ENCEPHALOMYOCARDITIS VIRUS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M16020; AAA47930.1; -.
CC PIR; A29535; GNNYTM.
CC PDB; 1TMF; 31-OCT-93.
CC MEROPS; C03.009; -.
CC MEROPS; U29.001; -.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR001676; RNV.
CC Pfam; PF00073; rhv; 3.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC KW Polyprotein; Coat protein; Core protein; Transferase;
CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
CC 3D-structure.
CC PROPEP 1 76
CC CHAIN 77 147
CC CHAIN 148 414
CC CHAIN 415 646
CC CHAIN 647 922
CC CHAIN 923 1064
CC CHAIN 1065 1191
CC CHAIN 1192 1517
CC CHAIN 1518 1605
CC CHAIN 1606 1625
CC CHAIN 1626 1842
CC CHAIN 1843 2303
CC LIPID 77 77
CC ACT_SITE 1793 1793
CC ACT_SITE 1811 1811
CC HELIX 103 106
CC HELIX 158 160
CC STRAND 162 166
CC TURN 167 168
CC STRAND 169 173
CC STRAND 180 181
CC STRAND 200 200
CC HELIX 204 206
CC STRAND 210 217
CC TURN 219 220
CC TURN 223 224
CC STRAND 225 226
CC STRAND 228 231
CC TURN 233 235
CC TURN 239 240
CC HELIX 241 247
CC STRAND 250 253
CC STRAND 256 262
CC TURN 267 268
CC STRAND 272 278
```

FT TURN 297 298
FT STRAND 305 305
FT TURN 313 314
FT TURN 325 326
FT TURN 330 334
FT STRAND 338 341
FT TURN 343 345
FT STRAND 348 353
FT STRAND 362 363
FT STRAND 364 366
FT STRAND 370 378
FT TURN 384 385
FT STRAND 390 399
FT STRAND 403 406
FT STRAND 428 429
FT TURN 430 431
FT TURN 459 463
FT STRAND 478 478
FT STRAND 483 483
FT STRAND 493 495
FT TURN 505 506
FT TURN 508 511
FT TURN 513 514
FT STRAND 518 520
FT TURN 524 529
FT TURN 533 534
FT STRAND 536 543
FT TURN 546 547
FT HELIX 554 557
FT TURN 558 559
FT STRAND 561 562
FT STRAND 565 566
FT STRAND 572 575
FT STRAND 586 587
FT STRAND 602 611
FT TURN 614 615
FT STRAND 620 620
FT STRAND 622 627
FT STRAND 633 635
FT TURN 652 653
FT TURN 680 684
FT STRAND 691 692
FT TURN 707 708
FT STRAND 713 714
FT TURN 760 764
FT STRAND 770 780
FT STRAND 791 794
FT TURN 796 797
FT STRAND 820 822
FT STRAND 831 835
FT STRAND 844 845
FT TURN 865 866
FT STRAND 871 874
FT STRAND 880 892
SQ SEQUENCE 2303 AA; 256280 MW; E2C7737DFDBEB786 CRC64;

Query Match 20.7%; Score 6; DB 1; Length 2303;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 LCKLPT 13
|||||
Db 461 LCKLPT 466

RESULT 25
POLG_TMEVG STANDARD; PRT: 2303 AA.
AC P08545; Q88593;
DT 01-AUG-1988 (Rel. 08, Created)

DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; P3A; Genome-linked protein VP3; Picornain 3C
DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D
DE (EC 2.7.7.48)].
OS Theiler's murine encephalomyelitis virus (strain GDVII) (TMEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Cardiovirus.
OX NCBI_TaxID=12127;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=88265847; PubMed=2838951;
RA Pevear D.C., Borkowski J., Calenoff M., Oh C.K., Ostrowski B.,
RA Lipton H.L.;
RT "Insights into Theiler's virus neurovirulence based on a genomic
RT comparison of the neurovirulent GDVII and less virulent Bean
RT strains.";
RL Virology 165:1-12(1988).
RN [2]
SEQUENCE OF 1704-2303 FROM N.A.
RX MEDLINE=87061197; PubMed=3023668;
RA Ozden S., Tangy F., Chamorro M., Brahic M.;
RT "Theiler's virus genome is closely related to that of
RT encephalomyocarditis virus, the prototype cardiovirus.";
RL J. Virol. 60:1163-1165(1986).
CC -!- FUNCTION: IT IS THOUGHT THAT THE P2C PROTEIN ATTACHES TO VESICULAR
CC MEMBRANES AND IS ASSOCIATED WITH VIRAL RNA SYNTHESIS.
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- SIMILARITY: CLOSELY RELATED TO ENCEPHALOMYOCARDITIS VIRUS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M20562; AAA47929.1; -;
CC DR EMBL; M14703; AAA47933.1; -;
CC DR PIR; A29193; GNNVTP.
CC DR PIR; A26100; A26100.
CC DR HSSP; Q88571; LTME.
CC DR MEROPS; C03.009; -;
CC DR MEROPS; U29.001; -;
CC DR InterPro; IPR000605; RNA_helicase.
CC DR InterPro; IPR001205; RNA_pol_P3D.
CC DR InterPro; IPR001676; Rhv.
CC DR Pfam; PF00073; rhv; 3.
CC DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC DR Pfam; PF00910; RNA_helicase; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate.
FT PROPEP 1 76 LEADER PEPTIDE.
FT CHAIN 77 147 COAT PROTEIN VP4 (P1A).
FT CHAIN 148 414 COAT PROTEIN VP2 (P1B).
FT CHAIN 415 646 COAT PROTEIN VP3 (P1C).
FT CHAIN 647 922 COAT PROTEIN VP1 (P1D).
FT CHAIN 923 1064 CORE PROTEIN P2A.
FT CHAIN 1065 1191 CORE PROTEIN P2B.
FT CHAIN 1192 1517 CORE PROTEIN P2C.
FT CHAIN 1518 1605 CORE PROTEIN P3A.
FT CHAIN 1605 1625 GENOME-LINKED PROTEIN VP3 (P3B).
FT CHAIN 1626 1842 PICORNAIN 3C.
FT CHAIN 1843 2303 RNA-DIRECTED RNA POLYMERASE P3D.
FT LIPID 77 77 MYRISTATE (BY SIMILARITY).

```
FT ACT_SITE 1793 1793 PROTEASE (POTENTIAL).
FT ACT_SITE 1811 1811 PROTEASE (POTENTIAL).
FT CONFLICT 1747 1750 VTGI -> CYRD (IN REF. 2).
FT CONFLICT 1992 2000 DDVVYQSFL -> GRRCPLPIIF (IN REF. 2).
FT CONFLICT 2003 2003 E -> Q (IN REF. 2).
FT CONFLICT 2008 2008 E -> H (IN REF. 2).
FT CONFLICT 2046 2046 F -> L (IN REF. 2).
FT CONFLICT 2128 2134 LIRGGLP -> YSMGPA (IN REF. 2).
SQ SEQUENCE 2303 AA; 256342 MW; 5D0FBE6E47F72A04 CRC64;

Query Match 20.7%; Score 6; DB 1; Length 2303;
Best Local Similarity 100.0%; Pred. No. 1;le+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LCKLPT 13
Db 461 LCKLPT 466

RESULT 26
LPTN_PROVU STANDARD; PRT; 34 AA.
ID LPTN_PROVU STANDARD; PRT; 34 AA.
AC P28779;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Tryptophanase leader peptide.
GN TNAL OR TNAC.
OS Proteus vulgaris.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Proteus.
OX NCBI_TaxID=585;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=93015658; PubMed=1400314;
RX Kamath A.V., Yanofsky C.;
RT "Characterization of the tryptophanase operon of Proteus vulgaris.
RT Cloning, nucleotide sequence, amino acid homology, and in vitro
RT synthesis of the leader peptide and regulatory analysis.";
RJ J. Biol. Chem. 267:19978-19985(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M93277; AAA25663.1; -.
CC PIR; A44038; A44038.
CC Leader peptide.
SQ SEQUENCE 34 AA; 4140 MW; 4C900172C9AD58F3 CRC64;

Query Match 17.2%; Score 5; DB 1; Length 34;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 DSELA 18
Db 24 DSELA 28

RESULT 27
VGK_BPAL3 STANDARD; PRT; 53 AA.
AC P31282;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE K protein.
```

```
GN K. Bacteriophage alpha-3.
OS Viruses; ssDNA viruses; Microviridae; Microvirus.
OX NCBI_TaxID=10849;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92223109; PubMed=1532908;
RA Kodaira K.-I., Nakano K., Okada S., Taketo A.;
RT "Nucleotide sequence of the genome of the bacteriophage alpha 3:
RT interrelationship of the genome structure and the gene products with
RT those of the phages, phi XI74, G4 and phi K.";
RL Biochim. Biophys. Acta 1130:277-288(1992).
CC -!- FUNCTION: NO FUNCTION HAS YET BEEN ASCRIBED TO GENE K PROTEIN.
CC -!- MISCELLANEOUS: GENE K OVERLAPS GENES B, A, AND C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X60322; CAA42876.1; -.
CC PIR; S22326; S22326.
SQ SEQUENCE 53 AA; 6235 MW; 2B502C9EAEBCD223 CRC64;

Query Match 17.2%; Score 5; DB 1; Length 53;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LLCKL 11
Db 44 LLCKL 48

RESULT 28
VGK_BPPHK STANDARD; PRT; 53 AA.
ID VGK_BPPHK STANDARD; PRT; 53 AA.
AC Q38038;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE K protein.
GN K.
OS Bacteriophage phi-K.
OX Viruses; ssDNA viruses; Microviridae; Microvirus.
OX NCBI_TaxID=10848;
RN [1]
RP SEQUENCE FROM N.A.
RA Kodaira K.-I., Tadokoro H., Suzuki K., Oki M., Manto S., Taketo A.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: NO FUNCTION HAS YET BEEN ASCRIBED TO GENE K PROTEIN.
CC -!- MISCELLANEOUS: GENE K OVERLAPS GENES B, A, AND C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X60323; CAA42887.1; -.
CC PIR; S22326; S22326.
SQ SEQUENCE 53 AA; 6207 MW; FF602C9EAEBCAB2 CRC64;

Query Match 17.2%; Score 5; DB 1; Length 53;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LLCKL 11
```

```
Db 44 LLCKL 48
|||||
RESULT 29
Y524_MYCLE
ID Y524_MYCLE STANDARD; PRT; 79 AA.
AC O05755;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 8.6 kDa protein ML2219.1.
GN ML2219.1 OR ML2219A OR MLCB5.24.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
[1]
SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC -!- SIMILARITY: BELONGS TO THE UPF0062 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL; Z95151; CAB08428.1; -.
DR EMBL; AL583924; CAB31175.1; -.
DR Leproma; ML2219A; -.
DR InterPro; IPR003850; UPF0062.
DR Pfam; PF02700; UPF0062.1
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 79 AA; 8646 MW; E8ICE41BEC0CB59B CRC64;

Query Match 17.2%; Score 5; DB 1; Length 79;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 DSELA 18
Db 53 DSELA 57
|||||
RESULT 30
Y65_BPT7
ID Y65_BPT7 STANDARD; PRT; 84 AA.
AC P03800;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-MAR-1989 (Rel. 10, Last annotation update)
DE Hypothetical gene 6.5 protein.
GN 6.5.
OS Bacteriophage T7.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like phages.
OX NCBI_TaxID=10760;

[1]
SEQUENCE FROM N.A.
RX MEDLINE=83241725; PubMed=6864790;
RA Dunn J.J., Studier F.W.;
RT "Complete nucleotide sequence of bacteriophage T7 DNA and the
RT locations of T7 genetic elements.";
RL J. Mol. Biol. 166:477-535(1983).
CC
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL; V01146; CAA24420.1; -.
DR PIR; A04425; Q6BP57.
DR PIR; S42318; S42318.
KW Hypothetical protein.
SQ SEQUENCE 84 AA; 9475 MW; 597AD49E02E24445 CRC64;

Query Match 17.2%; Score 5; DB 1; Length 84;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 INQLL 8
Db 5 INQLL 9
|||||
RESULT 31
RS20_BORBU
ID RS20_BORBU STANDARD; PRT; 85 AA.
AC P49394; O51249;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S20.
GN RPST OR RB0233.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
[1]
SEQUENCE FROM N.A.
RA Tilly K.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Winn M.,
RA Dougherty B., Tomb J.-P., Fleischmann R.D., Richardson D.,
RA Peterson J., Kervatage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Uterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
CC -!- FUNCTION: Binds directly to 16S ribosomal RNA (By similarity).
CC -!- SIMILARITY: BELONGS TO THE S20P FAMILY OF RIBOSOMAL PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
```

```
CC -----
DR EMBL; U35673; AAB41459.1; -.
DR EMBL; AE001133; AAC66616.1; ALT_INIT.
DR TIGR; BB0233; -.
DR InterPro; IPR002583; Ribosomal_S20p.
DR Pfam; PF01649; Ribosomal_S20p; 1.
DR ProDom; PD004231; Ribosomal_S20p; 1.
KW Ribosomal protein; rRNA-binding; Complete proteome.
SQ SEQUENCE 85 AA; 9947 MW; EBAC874CD12397FA CRC64;

Query Match 17.2%; Score 5; DB 1; Length 85;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRCIN 5
DB 33 KRCIN 37

RESULT 32
Y898_MYCTU STANDARD; PRT; 87 AA.
ID Y898_MYCTU
AC Q10566;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 9.9 kDa protein Rv0898c.
GN Rv0898c OR MT0921.1 OR MTCY31.26C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagsels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.B., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z73101; CAA97373.1; -.
DR EMBL; AE006979; AAK45168.1; -.
DR TIGR; MT0921.1; -.
DR TubercuList; Rv0898c; -.
KW Hypothetical protein; Complete proteome.
```

```
SQ SEQUENCE 87 AA; 9920 MW; 79D055AB95C09BB2 CRC64;

Query Match 17.2%; Score 5; DB 1; Length 87;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSE 16
DB 7 PTDSE 11

RESULT 33
RL36_DAUCA STANDARD; PRT; 106 AA.
ID RL36_DAUCA
AC P52866;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE 60S ribosomal protein L36.
GN RPL36.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. DANVERS HALF-LONG;
RA Lin X., Hwang G.J., Zimmerman J.L.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE L36E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U47095; AAB01095.1; ALT_INIT.
DR InterPro; IPR000509; Ribosomal_L36e.
DR Pfam; PF01158; Ribosomal_L36e; 1.
DR ProDom; PD009192; Ribosomal_L36e; 1.
DR ProSITE; PS01190; RIBOSOMAL_L36E; 1.
KW Ribosomal protein.
SQ SEQUENCE 106 AA; 12250 MW; 016BAFFC18409B24 CRC64;

Query Match 17.2%; Score 5; DB 1; Length 106;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELAPR 20
DB 24 ELAPR 28

RESULT 34
PSC2_RAT STANDARD; PRT; 112 AA.
ID PSC2_RAT
AC P02781;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Prostetic steroid-binding protein C2 chain precursor (Prostatein
DE peptide C2).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
```

SEQUENCE FROM N.A.
 MEDLINE=871146484; PubMed=2881277;
 Delaey B., Dirckx L., Decourt J.-L., Claessens F., Peeters B.,
 Rombauts W.;
 "Rat prostatic binding protein: the complete sequence of the C2 gene
 and its flanking regions.";
 Nucleic Acids Res. 15:1627-1641(1987).
 [2]
 SEQUENCE OF 21-112.
 RP MEDLINE=83209619; PubMed=6343081;
 Peeters B., Heyns W., Mous J., Rombauts W.;
 "Structural studies on rat prostatic binding protein. The primary
 structure of component C2 from subunit S.";
 Eur. J. Biochem. 132:669-679(1983).
 [3]
 SEQUENCE OF 1-100 FROM N.A.
 RP MEDLINE=82220075; PubMed=6896362;
 Parker M., Needham M., White R.;
 "Prostatic steroid binding protein: gene duplication and steroid
 binding.";
 Nature 298:92-94(1982).
 CC -!- FUNCTION: PART OF PROSTATEIN WHICH IS THE MAJOR SECRETORY
 CC GLYCOPROTEIN OF VENTRAL PROSTATE GLAND.
 CC -!- SUBUNIT: PROSTATEIN IS COMPOSED OF THREE DIFFERENT PEPTIDES CALLED
 CC C1, C2 AND C3. THESE FORM COVALENT C1:C3 (F) AND C2:C3 (S)
 CC HETERODIMERS WHOSE NONCOVALENT ASSOCIATION FORMS TETRAMERIC
 CC (C1:C3/C3:C2) PROSTATEIN MOLECULES.
 CC -!- SUBCELLULAR LOCATION: Secreted
 CC -!- LINKED BY THREE DISULFIDE BONDS TO C3.
 CC -!- MISCELLANEOUS: THE HETERODIMER CAN BIND NON-POLAR STEROIDS,
 CC CHOLESTEROL AND A GROUP OF SMALL PROLINE-RICH PEPTIDES.
 CC -!- SIMILARITY: BELONGS TO THE UTEROGLOBIN FAMILY. LIPOPHILIN
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X05034; CAA28708.1; -
 CC EMBL; V01256; CAA24569.1; -
 CC EMBL; J00776; AAA51641.1; -
 CC PIR; A03251; B0RT2
 CC PIR; A26671; A26671.
 CC InterPro; IPR003628; Uteroglobn_sub.
 CC InterPro; IPR000329; Uterogloblin.
 CC ProDom; PD012475; Uteroglobn_sub; 1.
 CC PROSITE; PS00403; UTEROGLOBIN_1; FALSE NEG.
 CC PROSITE; PS00404; UTEROGLOBIN_2; FALSE NEG.
 CC Steroid-binding; Signal.
 CC SIGNAL 1 20
 CC CHAIN 21 112 PROSTATIC STEROID-BINDING PROTEIN C2
 CC CHAIN.
 CC MOD_RES 21 21
 CC DISULFID 28 28 INTERCHAIN (WITH C3) (PROBABLE).
 CC DISULFID 69 69 INTERCHAIN (WITH C3) (PROBABLE).
 CC DISULFID 92 92 INTERCHAIN (WITH C3) (PROBABLE).
 CC CONFLICT 26 26 MISSING (IN REF. 3).
 CC CONFLICT 88 88 I -> T (IN REF. 3).
 CC CONFLICT 96 112 VWLQNFPRGRWFSEIN -> YGYK (IN REF. 3).
 CC SEQUENCE 112 AA; 12828 MW; DA65A6A82E677864 CRC64;

Query Match 17.2%; Score 5; DB 1; Length 112;
 Best Local Similarity 100.0%; Pred. No. 90;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KRCIN 5
 Db 67 KRCIN 71

RESULT 35
 R36B_ARATH
 ID R36B_ARATH STANDARD; PRT; 112 AA.
 AC Q9M352;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE 60S ribosomal protein L36-2
 GN RPL36B OR AT3G53740 OR P5K20_40.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 [1]
 SEQUENCE FROM N.A.
 RP STRAIN=CV. COLUMBIA;
 RX MEDLINE=21016720; PubMed=11130713;
 RA Salanoubat M., Lemcke K., Rieger M., Ansoorge W., Unseld M.,
 RA Farman B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,
 RA Delzeny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
 RA De Simone V., Choise N., Artiguenave F., Robert C., Brotier P.,
 RA Wincker P., Cattolico L., Weissbach J., Saurin W., Quetier F.,
 RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
 RA Wurbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,
 RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
 RA Vezzi A., D'Angelo M., Fallavincini A., Toppo S., Simonati B.,
 RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G.,
 RA Reichelt J., Scharfe M., Schoen O., Barges M., Terol J., Climent J.,
 RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
 RA Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D.,
 RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
 RA Monfort A., Argirou A., Flores M., Liguori R., Vitale D.,
 RA Manhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
 RA Mayer K.F.K., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
 RA Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,
 RA Cressy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
 RA Pai G., Mlitscher J., Sallers P., Gill J.E., Feldblyum T.V.,
 RA Preuss D., Lin X., Niernan W.C., Salzberg S.L., White O., Venter J.C.,
 RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
 RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
 RA Watanabe A., Yamada M., Yasuda M., Tabata S.;
 "Sequence and analysis of chromosome 3 of the plant Arabidopsis
 thaliana";
 RL Nature 408:820-822(2000).
 CC -!- SIMILARITY: BELONGS TO THE L36E FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AL132960; CAB88336.1; -
 CC InterPro; IPR000509; Ribosomal_L36e.
 CC Pfam; PF01158; Ribosomal_L36e; 1.
 CC ProDom; PD009192; Ribosomal_L36e; 1.
 CC PROSITE; PS01190; RIBOSOMAL_L36e; 1.
 CC Ribosomal protein; Multigene family.
 CC SEQUENCE 112 AA; 12676 MW; 3485A3D86A63C423 CRC64;

Query Match 17.2%; Score 5; DB 1; Length 112;
 Best Local Similarity 100.0%; Pred. No. 90;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 16 ELAPR 20

Db 24 ELAPR 28
|||||

```
RESULT 36
R36A_ARATH STANDARD; PRT; 113 AA.
AC R36A_ARATH
AC R36A_ARATH
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE 60S ribosomal protein L36-1.
GN RPL36A OR AT2G37600 OR F13M22.10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucrods II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA Moffat K.S., Cronin L.A., Shen M.D., Pai G., Van Aken S., Umayam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana."
RN Nature 402:761-768(1999).
RP [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Shinozaki A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RC Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE L36E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

Query Match 17.2%; Score 5; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 ELAPR 20
|||||

Db 24 ELAPR 28

```
RESULT 37
WN15_HUMAN STANDARD; PRT; 120 AA.
ID WN15_HUMAN
AC O14905;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE WNT-15 protein (Fragment).
GN WN15.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98110581; PubMed=9441749;
RA Bergstein I., Eisenberg L.M., Bhallerao J., Jenkins N.A.,
RA Copeland N.G., Osborne M.P., Bowcock A.M., Brown A.M.C.;
RT "Isolation of two novel WNT genes, WNT14 and WNT15, one of which
RT (WNT15) is closely linked to WNT3 on human chromosome 17q21."
RL Genomics 46:450-458(1997).
CC -!- FUNCTION: LIGAND FOR MEMBERS OF THE FRIZZLED FAMILY OF SEVEN
CC TRANSMEMBRANE RECEPTORS. PROBABLE DEVELOPMENTAL PROTEIN. MAY BE A
CC SIGNALING MOLECULE WHICH AFFECT THE DEVELOPMENT OF DISCRETE
CC REGIONS OF TISSUES. IS LIKELY TO SIGNAL OVER ONLY FEW CELL
CC DIAMETERS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Possibly secreted and associates with the
CC extracellular matrix.
CC -!- SIMILARITY: BELONGS TO THE WNT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

Query Match 17.2%; Score 5; DB 1; Length 120;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LAPRS 21
|||||

Db 59 LAPRS 63

```
RESULT 38
Y055_MYCGE STANDARD; PRT; 123 AA.
ID Y055_MYCGE
AC P47301; O49349;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MG055.
GN MG055.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
```

```

RN  SEQUENCE FROM N.A.
RC  STRAIN=ATCC 33530 / G-37;
RX  MEDLINE=96026346; PubMed=7569993;
RA  Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA  Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA  Fritchman J.B., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA  Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA  Tomb J.-F., Dougherty B.O., Bott K.F., Hu P.-C., Lucier T.S.,
RA  Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT  "The minimal gene complement of Mycoplasma genitalium.";
RL  Science 270:397-403(1995).
RN  [2]
RP  SEQUENCE OF 1-108 FROM N.A.
RC  STRAIN=ATCC 33530 / G-37;
RX  MEDLINE=94075230; PubMed=8253680;
RA  Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
RT  "A survey of the Mycoplasma genitalium genome by using random
RT  sequencing.";
RL  J. Bacteriol. 175:7918-7930(1993).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; U39684; AAC71271.1; -;
DR  EMBL; U02240; AAA03395.1; ALT_INIT.
DR  TIGR; MG055; -.
KW  Hypothetical protein; Transmembrane; Complete proteome.
FT  TRANSMEM 91 111 POTENTIAL.
SQ  SEQUENCE 123 AA; 14460 MW; 3EC4967DC0FBC30 CRC64;

Query Match 17.2%; Score 5; DB 1; Length 123;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 INQLL 8
Db 111 INQLL 115

RESULT 39
YJS4_YEAST
ID YJS4_YEAST STANDARD; PRT; 123 AA.
AC P46984;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical 13.6 kDa protein in SWEL-ATP12 intergenic region.
GN YJL184W OR J0420.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Obermaier B., Piravandi E., Rinke M., Domdey H.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z49459; CAA89479.1; -.

```

```

DR SGD; S0003720; YJL184W.
KW Hypothetical protein.
SQ SEQUENCE 123 AA; 13605 MW; 437EAF6CF1332CC8 CRC64;

Query Match 17.2%; Score 5; DB 1; Length 123;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 INQLL 8
Db 109 INQLL 113

RESULT 40
LECC_POLMI
ID LECC_POLMI STANDARD; PRT; 125 AA.
AC P16108;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lectin.
OS Polyandrocarpa misakiensis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Styelidae; Polyandrocarpa.
OX NCBI_TaxID=7723;
RN [1]
RP SEQUENCE.
RX MEDLINE=90110175; PubMed=2295629;
RA Suzuki T., Takagi T., Furukohri T., Kawamura K., Nakauchi M.;
RT "A calcium-dependent galactose-binding lectin from the tunicate
RT Polyandrocarpa misakiensis. Isolation, characterization, and amino
RT acid sequence.";
RL J. Biol. Chem. 265:1274-1281(1990).
CC -1- FUNCTION: ROLE IN THE DEFENCE SYSTEM OF THE ORGANISM AGAINST
CC -1- MICROORGANISMS. THIS CALCIUM-BINDING LECTIN BINDS GALACTOSE.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
DR InterPro; IPR001304; lectin_c.
DR Pfam; PF000059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
DR PROSITE; PSS0041; C-TYPE_LECTIN_2; 1.
KW Calcium; Lectin.
FT DOMAIN 19 121 C-TYPE LECTIN (SHORT FORM).
FT DISULFID 21 119
FT DISULFID 96 111
FT SEQUENCE 125 AA; 14033 MW; 5C9B98F6F8B00C4A CRC64;

Query Match 17.2%; Score 5; DB 1; Length 125;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 LPTDS 15
Db 75 LPTDS 79

RESULT 41
RT12_PROWI
ID RT12_PROWI STANDARD; PRT; 125 AA.
AC P46744;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Mitochondrial ribosomal protein S12.
GN RPS12.
OS Prototheca wickerhamii.
OC Mitochondrion.
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC Chlorellaceae; Prototheca.
OX NCBI_TaxID=3111;

```

```

[1]
RN  SEQUENCE FROM N.A.
RC  STRAIN-263-11;
RA  MEDLINE-94180393; PubMed-8133522;
RW  Wolff G., Plante I., Lang B.F., Kueck U., Burger G.;
RT  "Complete sequence of the mitochondrial DNA of the chlorophyte alga
RL  Prototheca wickerhamii. Gene content and genome organization.";
RJ  J. Mol. Biol. 237:75-86(1994).
CC  -1- SUBCELLULAR LOCATION: Mitochondrial.
CC  -1- SIMILARITY: BELONGS TO THE S12P FAMILY OF RIBOSOMAL PROTEINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; U02970; AAD12660.1; -.
DR  Mende; 12225; PROWL; rps12.1.
DR  InterPro; IPR000230; Ribosomal_S12.
DR  Pfam; PF00164; Ribosomal_S12; 1.
DR  PRINTS; PR01034; RIBOSOMAL_S12.
DR  ProDom; PD000576; Ribosomal_S12; 1.
DR  PROSITE; PS00055; RIBOSOMAL_S12; 1.
KW  Ribosomal protein; Mitochondrion.
SQ  SEQUENCE 125 AA; 13638 MW; FFB08F9191B76880 CRC64;

Query Match      17.2%; Score 5; DB 1; Length 125;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  4 INQLL 8
DB  4 INQLL 8

RESULT 42
RL17_XANCP
ID  RL17_XANCP STANDARD; PRT; 127 AA.
AC  Q923E6;
DT  30-MAY-2000 (Rel. 39, Created)
DT  30-MAY-2000 (Rel. 39, Last sequence update)
DE  30-MAY-2000 (Rel. 39, Last annotation update)
DE  50S ribosomal protein L17.
GN  RPLQ.
OS  Xanthomonas campestris (pv. campestris).
OC  Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC  Xanthomonas.
OX  NCBI_TaxID=340;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-XC11;
RA  Huang C.F., Yang M.T.;
RL  Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC  -1- SIMILARITY: BELONGS TO THE L17P FAMILY OF RIBOSOMAL PROTEINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; U79735; AAD00326.1; -.
DR  InterPro; IPR000456; Ribosomal_L17.
DR  Pfam; PF01196; Ribosomal_L17; 1.
DR  ProDom; PD004277; Ribosomal_L17; 1.
DR  PROSITE; PS01167; RIBOSOMAL_L17; 1.
KW  Ribosomal protein.

```

```

SQ  SEQUENCE 127 AA; 14347 MW; B56CD1E2B664C7DF CRC64;

Query Match      17.2%; Score 5; DB 1; Length 127;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  16 ELAPR 20
DB  82 ELAPR 86

RESULT 43
Y05Q_BPT4
ID  Y05Q_BPT4 STANDARD; PRT; 128 AA.
AC  P13303; P39235; Q96212;
DT  01-JAN-1990 (Rel. 13, Created)
DT  15-DEC-1998 (Rel. 37, Last sequence update)
DT  01-MAR-2002 (Rel. 41, Last annotation update)
DE  Hypothetical 14.6 kDa protein in mobD-ri intergenic region.
GN  Y05O OR MOBD.7 OR TK.-3 OR RI.-1 OR 58.3.
OS  Bacteriophage T4.
OC  Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC  T4-like phages.
OX  NCBI_TaxID=10665;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Mzhavia N., Marusich E., Djavakhishvili T., Neitzel J., Peterson S.,
RA  Awaysa M., Eidermiller J., Canada D., Tracy J., Gaillbreath K.,
RA  Paddison P., Anderson B., Stidham T., Blattner F., Kutter E.M.;
RT  "The 10.7 kb 'nonessential' region of bacteriophage T4 between the
RT  genes tk and nrdc: twenty new t4 genes, generally conserved among
RT  T-even phages.";
RL  Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RA  Kutter E., Arisaka F., Kunisawa T., Tsugita A., Mosig G.,
RA  Mesyanzhinov V., Ruger W., Stidham T., Thomas E.;
RT  "Bacteriophage T4 genome analysis.";
RL  Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN  [3]
RP  SEQUENCE OF 1-27 FROM N.A.
RX  MEDLINE-87066735; PubMed-3024113;
RA  Valerie K., Stevens J., Lynch M., Henderson E.E., de Riel J.K.;
RT  "Nucleotide sequence and analysis of the 58.3 to 65.5-kb early region
RT  of bacteriophage T4.";
RL  Nucleic Acids Res. 14:8637-8654(1986).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; U76612; AAB26963.1; -.
DR  EMBL; AF158101; AAD42598.1; -.
DR  EMBL; X04567; CAA28237.1; -.
KW  Hypothetical protein.
SQ  SEQUENCE 128 AA; 14647 MW; FEF8D7A8D96A4B68 CRC64;

Query Match      17.2%; Score 5; DB 1; Length 128;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  10 KLPTD 14
DB  80 KLPTD 84

RESULT 44

```

```
VNSL_BRSVA
ID VNSL_BRSVA STANDARD; PRT; 136 AA.
AC Q65694;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Nonstructural protein 1 (Nonstructural protein 1C).
GN 1C OR NS1.
OS Bovine respiratory syncytial virus (strain A51908) (BRS).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
ON NCBI_TaxID=11247;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=95146950; PubMed=7844532;
RA Pastey M.K., Samal S.K.;
RT "Nucleotide sequence analysis of the non-structural NS1 (1C) and NS2
RL J. Gen. Virol. 76:193-197(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U15937; AAA5671.1; -
CC Nonstructural protein.
KW Nonstructural protein.
SQ SEQUENCE 136 AA; 15246 MW; AEB6A89D7BED2D24 CRC64;

Query Match 17.2%; Score 5; DB 1; Length 136;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 DSELA 18
   |||||
Db 119 DSELA 123

RESULT 45
YE98_HAEIN
ID YE98_HAEIN STANDARD; PRT; 139 AA.
AC P44222;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein H11498.
GN H11498.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-P., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd."
RL Science 269:496-512(1995).
RN [2]
RP IDENTIFICATION BY MASS SPECTROMETRY.

RX MEDLINE=20137488; PubMed=10675023;
RA Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,
RA Gray C., Fountoulakis M.;
RT "Two-dimensional map of the proteome of Haemophilus influenzae.";
RL Electrophoresis 21:411-429(2000).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32826; AAC23148.1; -
CC TIGR; H11498; -
KW Transmembrane; Complete proteome.
FT TRANSMEM 43 59 POTENTIAL.
SQ SEQUENCE 139 AA; 15625 MW; EADE7EADA58AA7F CRC64;

Query Match 17.2%; Score 5; DB 1; Length 139;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LCKLP 12
   |||||
Db 19 LCKLP 23

RESULT 46
NDK_AQUAE
ID NDK_AQUAE STANDARD; PRT; 142 AA.
AC O67528;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nucleoside diphosphate kinase (EC 2.7.4.6) (NDK) (NDP kinase)
DE (Nucleoside-2-P kinase).
GN NDK OR AQ.1590.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358(1998).
CC -1- FUNCTION: MAJOR ROLE IN THE SYNTHESIS OF NUCLEOSIDE TRIPHOSPHATES
CC OTHER THAN ATP.
CC -1- CATALYTIC ACTIVITY: ATP + nucleoside diphosphate = ADP +
CC nucleoside triphosphate.
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE NDK FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000746; AAC07481.1; -
CC HSSP; P15266; INHK.
CC InterPro: IPR001564; NDK.
CC Pfam; PF00334; NDK; 1.
```

```

DR ProDom; PD001018; NDK; 1.
KW PROSITE; PS00469; NDP_KINASES; 1.
KW Transferase; Kinase; ATP-binding; Complete proteome.
FT ACT_SITE 120 120 BY SIMILARITY.
SQ SEQUENCE 142 AA; 15942 MW; 189B5DFA0317E998 CRC64;

Query Match 17.2%; Score 5; DB 1; Length 142;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSE 16
Db 92 PTDSE 96

RESULT 47
UP03_ECOLI
ID UP03_ECOLI STANDARD; PRT; 144 AA.
AC P37903; P76074; P76853;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Unknown protein 2D-000B3L from 2D-page.
GN YNAF OR B1376.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
RA Sampel G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
RN [3]
RP PARTIAL SEQUENCE OF 1-12.
RC STRAIN=K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
in the genome of Escherichia coli K-12.";
RL Electrophoresis 18:1259-1313(1997).
RN [4]
RP SEQUENCE OF 1-11.
RC STRAIN=K12 / W3110;
RA Pasquali C., Sanchez J.-C., Ravier F., Golaz O., Hughes G.J.,
RA Frutiger S., Paquet N., Wilkins M., Appel R.D., Bairoch A.,
RA Hochstrasser D.F.;
RL Submitted (SEP-1994) to the SWISS-PROT data bank.
CC -!- SIMILARITY: BELONGS TO THE UPF0022 (USPA) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; AE000234; AAC74458.1; ALT_INIT.
DR EMBL; D90775; BAA14980.1; -.
DR SWISS-2DPAGE; P37903; COLI.
DR EcoGene; EG12674; ynaF.
DR InterPro; IPR000041; Usp.
DR Pfam; PF00582; Usp; 1.
KW Complete proteome.
SQ SEQUENCE 144 AA; 16016 MW; F30D0FC681EBF32D CRC64;

Query Match 17.2%; Score 5; DB 1; Length 144;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KLPTD 14
Db 80 KLPTD 84

RESULT 48
YMBW_YEAST
ID YMBW_YEAST STANDARD; PRT; 150 AA.
AC Q03579;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 17.2 kDa protein in PRC1-ADE4 intergenic region.
GN YMR298W.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Badcock K., Churcher C., Barrell B.G., Rajandream M.A.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; X80836; CAA56807.1; -.
DR SGD; S0004913; YMR298W.
KW Hypothetical protein.
SQ SEQUENCE 150 AA; 17207 MW; 75DBA35225C3065C CRC64;

Query Match 17.2%; Score 5; DB 1; Length 150;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDSEL 17
Db 133 TDSEL 137

RESULT 49
YMEC_METEX
ID YMEC_METEX STANDARD; PRT; 153 AA.
AC Q49117;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical oxidoreductase in MEAA 3' region (EC 1.-.-.-) (ORFC)

```

```

DE (Fragment).
OS Methylobacterium extorquens.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
OX NCBI_TaxID=408;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AM1 / NCIMB 9133;
RX MEDLINE=97022083; PubMed=8868443;
RA Smith L.M., Meijer W.G., Dijkhuizen L., Goodwin P.M.;
RT "A protein having similarity with methylmalonyl-CoA mutase is
RT required for the assimilation of methanol and ethanol by
RT Methylobacterium extorquens AM1.";
RL Microbiology 142:675-684(1996).
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U28335; AAC44089.1; -.
DR HSP; Q12634; IYBV.
DR InterPro; IPR002198; ADH_short.
DR Pfam; PF00106; adh_short; 1.
DR PROSITE; PS00061; ADH_SHORT; 1.
KW Hypothetical protein; Oxidoreductase.
FT NON_TER 1
FT ACT_SITE 59 59 BY SIMILARITY.
SQ SEQUENCE 153 AA; 15745 MW; 628F9CD4A28DC70E CRC64;

Query Match 17.2%; Score 5; DB 1; Length 153;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 ELAPR 20
Db 76 ELAPR 80
|||||

RESULT 50
NUDH_PSEAE STANDARD; PRT; 159 AA.
AC Q9XAP2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable (d)inucleoside polyphosphate hydrolase (EC 3.6.1.-) (Ap4A
DE pyrophosphatase).
GN NUDH OR PA0336.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PA14;
RX MEDLINE=99162618; PubMed=10051655;
RA Tan M.-W., Rahme L.G., Sternberg J.A., Tompkins R.G., Ausubel F.M.;
RT "Pseudomonas aeruginosa killing of Caenorhabditis elegans used to
RT identify P. aeruginosa virulence factors.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:2408-2413(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,

```

```

RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -1- FUNCTION: Hydrolyzes diadenosine polyphosphate (By similarity).
CC -1- COFACTOR: BINDS TWO DIVALENT IONS: ZINC, MAGNESIUM, OR MANGANESE
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUDIX HYDROLASE FAMILY. NUDH SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF116285; AAD22458.1; -.
DR EMBL; AE004471; AAG03725.1; -.
DR HSP; P08337; ITUM.
DR InterPro; IPR000086; NUDIX_hydrolase.
DR Pfam; PF00293; NUDIX; 1.
DR PRINTS; PR00502; NUDIXFAMILY.
DR PROSITE; PS00893; NUDIX; 1.
KW Hydrolase; Magnesium; Manganese; Zinc; Complete proteome.
FT DOMAIN 38 59 NUDIX BOX.
SQ SEQUENCE 159 AA; 18758 MW; 05E119F7B82FCBE8 CRC64;

Query Match 17.2%; Score 5; DB 1; Length 159;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 ELAPR 20
Db 150 ELAPR 154
|||||

Search completed: August 22, 2002, 11:19:39
Job time: 166 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 11:16:53 ; Search time 42.29 Seconds
(without alignments)
118.630 Million cell updates/sec

Title: US-09-856-199-1

Perfect score: 29
Sequence: 1 KRCINQLCKLPDSELPASXCHCHNL 29

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1200 summaries

Database :

SPREMBL_19:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.todent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31.0	380	11	Q9Z2T7	Q9Z2T7 mus musculus
2	31.0	386	6	Q9N2H6	Q9N2H6 bos taurus
3	24.1	28	12	Q9E1T3	Q9E1T3 rabies viru
4	24.1	495	3	Q06858	Q06858 saccharomyc
5	24.1	1009	10	Q9M295	Q9M295 arabidopsis
6	20.7	98	10	Q42306	Q42306 arabidopsis
7	20.7	103	11	Q9JUF2	Q9JUF2 mus musculu
8	20.7	107	10	Q42046	Q42046 arabidopsis
9	20.7	125	10	Q94JW2	Q94JW2 arabidopsis
10	20.7	129	16	Q32096	Q32096 bacillus su
11	20.7	135	15	Q36891	Q36891 human immun
12	20.7	167	3	Q9Y874	Q9Y874 neurospora
13	20.7	180	10	Q9FP16	Q9FP16 arabidopsis
14	20.7	180	10	Q9SAV4	Q9SAV4 arabidopsis
15	20.7	207	2	Q9X903	Q9X903 streptomyce
16	20.7	209	2	Q9L0K0	Q9L0K0 streptomyce

17	20.7	212	10	Q9FGE8	Q9FGE8 arabidopsis
18	20.7	228	5	Q18836	Q18836 caenorhabdt
19	20.7	246	8	Q99882	Q99882 agarcicus su
20	20.7	246	8	Q99883	Q99883 agarcicus cr
21	20.7	246	8	Q9TFX8	Q9TFX8 agarcicus bl
22	20.7	248	2	Q9XDW2	Q9XDW2 rhodopseudo
23	20.7	259	16	Q931A2	Q931A2 rhizobium m
24	20.7	276	11	Q9DAU1	Q9DAU1 mus musculu
25	20.7	281	16	Q9KPY3	Q9KPY3 vibrio chol
26	20.7	295	16	Q9Z2PN5	Q9Z2PN5 rhizobium m
27	20.7	300	2	Q54221	Q54221 staphylococ
28	20.7	303	16	Q06279	Q06279 mycobacteri
29	20.7	308	11	Q35100	Q35100 rattus ratt
30	20.7	313	12	Q91TNS	Q91TNS tupiaa herp
31	20.7	346	16	P95153	P95153 mycobacteri
32	20.7	359	10	Q9LRX6	Q9LRX6 arabidopsis
33	20.7	364	16	Q9RZB3	Q9RZB3 deinococcus
34	20.7	391	6	Q9N0E8	Q9N0E8 tursiops tr
35	20.7	392	6	Q95171	Q95171 bison bison
36	20.7	396	2	Q06038	Q06038 lactococcus
37	20.7	404	2	Q9XSR5	Q9XSR5 streptomyce
38	20.7	412	10	Q9SG76	Q9SG76 arabidopsis
39	20.7	438	2	Q54823	Q54823 streptomyce
40	20.7	441	10	Q9AV62	Q9AV62 oryza sativ
41	20.7	442	10	Q9ZS99	Q9ZS99 arabidopsis
42	20.7	461	10	Q9M100	Q9M100 arabidopsis
43	20.7	499	3	Q96T44	Q96T44 aspergillus
44	20.7	503	3	Q9VL48	Q9VL48 drosophila
45	20.7	504	2	Q93A34	Q93A34 listeria mo
46	20.7	504	16	Q9ZBT3	Q9ZBT3 listeria mo
47	20.7	521	16	Q98A10	Q98A10 rhizobium l
48	20.7	602	5	Q9YUJ3	Q9YUJ3 drosophila
49	20.7	683	5	Q22187	Q22187 caenorhabdt
50	20.7	704	5	Q24213	Q24213 drosophila
51	20.7	704	5	Q24190	Q24190 drosophila
52	20.7	704	5	Q9V113	Q9V113 drosophila
53	20.7	711	4	Q9HC22	Q9HC22 homo sapien
54	20.7	711	4	Q9H4F2	Q9H4F2 homo sapien
55	20.7	711	4	Q9B1J1	Q9B1J1 homo sapien
56	20.7	734	5	Q9VTA9	Q9VTA9 drosophila
57	20.7	750	4	Q9Y5D9	Q9Y5D9 homo sapien
58	20.7	776	10	Q9SU76	Q9SU76 arabidopsis
59	20.7	813	4	Q9Y5D2	Q9Y5D2 homo sapien
60	20.7	816	5	Q9W0L6	Q9W0L6 drosophila
61	20.7	817	4	Q9Y5D0	Q9Y5D0 homo sapien
62	20.7	818	4	Q9Y5D1	Q9Y5D1 homo sapien
63	20.7	820	4	Q15039	Q15039 homo sapien
64	20.7	820	4	Q9Y5D7	Q9Y5D7 homo sapien
65	20.7	820	4	Q9Y5D3	Q9Y5D3 homo sapien
66	20.7	823	4	Q9Y5D5	Q9Y5D5 homo sapien
67	20.7	828	4	Q9Y5C9	Q9Y5C9 homo sapien
68	20.7	829	4	Q9Y5D4	Q9Y5D4 homo sapien
69	20.7	830	5	Q9VCW2	Q9VCW2 drosophila
70	20.7	837	4	Q9Y5D8	Q9Y5D8 homo sapien
71	20.7	850	4	Q9Y5E0	Q9Y5E0 homo sapien
72	20.7	882	12	Q88487	Q88487 theiler's e
73	20.7	890	5	Q45587	Q45587 caenorhabdt
74	20.7	920	12	Q88496	Q88496 theiler's e
75	20.7	920	12	Q88497	Q88497 theiler's e
76	20.7	922	12	Q56165	Q56165 theiler's e
77	20.7	922	12	Q88495	Q88495 theiler's e
78	20.7	922	12	Q88498	Q88498 theiler's e
79	20.7	928	11	Q08963	Q08963 rattus ratt
80	20.7	928	11	Q9Z1B0	Q9Z1B0 mus musculu
81	20.7	930	11	Q91XY5	Q91XY5 mus musculu
82	20.7	930	11	Q91XY9	Q91XY9 mus musculu
83	20.7	930	12	Q02472	Q02472 theiler's e
84	20.7	931	4	Q9Y5G9	Q9Y5G9 homo sapien
85	20.7	931	4	Q9Y5G8	Q9Y5G8 homo sapien
86	20.7	931	11	Q91XZ0	Q91XZ0 mus musculu
87	20.7	931	11	Q91XY6	Q91XY6 mus musculu
88	20.7	932	4	Q60330	Q60330 homo sapien
89	20.7	932	4	Q9Y5H1	Q9Y5H1 homo sapien

90	6	20.7	932	4	09Y5H0	09Y5h0 homo sapien	163	5	17.2	123	12	09W8B6	09W8B6 human cytom
91	6	20.7	932	4	09Y5G7	09Y5g7 homo sapien	164	5	17.2	124	4	09HDB6	09HDB6 homo sapien
92	6	20.7	932	4	09Y5G6	09Y5g6 homo sapien	165	5	17.2	125	15	073333	073333 human immun
93	6	20.7	932	4	09Y5G5	09Y5g5 homo sapien	166	5	17.2	126	11	09CUJ7	09CUJ7 mus musculus
94	6	20.7	932	4	09Y5G4	09Y5g4 homo sapien	167	5	17.2	126	16	09KPB6	09KPB6 vibrio chol
95	6	20.7	932	11	091XY3	091xy3 mus musculus	168	5	17.2	128	16	0980I4	0980I4 mycoplasma
96	6	20.7	932	11	091XY3	091xy3 mus musculus	169	5	17.2	129	8	025822	025822 plasmidom
97	6	20.7	932	11	091XY2	091xy2 mus musculus	170	5	17.2	131	8	021239	021239 reclinomona
98	6	20.7	932	11	091XY0	091xy0 mus musculus	171	5	17.2	132	16	091IX0	091IX0 pseudomonas
99	6	20.7	932	11	091XN9	091xn9 mus musculus	172	5	17.2	134	16	092SM1	092SM1 rhizobium m
100	6	20.7	933	11	091XN8	091xn8 mus musculus	173	5	17.2	135	16	P74399	P74399 synechocyst
101	6	20.7	935	4	09Y5H2	09Y5h2 homo sapien	174	5	17.2	136	4	09UGZ2	09UGZ2 homo sapien
102	6	20.7	935	11	091XY1	091xy1 mus musculus	175	5	17.2	136	12	09YJB7	09YJB7 bovine resp
103	6	20.7	936	4	09Y5H3	09Y5h3 homo sapien	176	5	17.2	136	12	091B60	091B60 bovine resp
104	6	20.7	1151	8	09XMH7	09Xmh7 neurospora	177	5	17.2	136	12	09YWM7	09YWM7 bovine resp
105	6	20.7	1319	16	055359	055359 synechocyst	178	5	17.2	136	12	09YWM7	09YWM7 bovine resp
106	6	20.7	1333	5	09VINO	09vino drosophila	179	5	17.2	138	3	09YWM5	09YWM5 schizosacch
107	6	20.7	1779	5	09VW49	09Vw49 drosophila	180	5	17.2	138	15	090277	090277 chimpanzee
108	6	20.7	2159	5	09Y8A6	09Y8a6 nodulispori	181	5	17.2	139	16	098F74	098F74 rhizobium l
109	6	20.7	2303	12	088595	088595 thelier's m	182	5	17.2	139	17	098F74	098F74 pyrococcus
110	6	20.7	2972	5	P90891	P90891 caenorhabdi	183	5	17.2	139	17	09V0G7	09V0G7 pyrococcus
111	5	17.2	29	11	063033	063033 rattus norv	184	5	17.2	140	16	092Bj8	092Bj8 listeria in
112	5	17.2	35	10	09S9G6	09S9g6 cucurbita (185	5	17.2	142	5	09VEI5	09VEI5 drosophila
113	5	17.2	47	5	09SRY0	09Sry0 drosophila	186	5	17.2	143	5	09YB36	09YB36 drosophila
114	5	17.2	49	16	098J53	098j53 rhizobium l	187	5	17.2	143	12	09ENM4	09ENM4 amsacta moo
115	5	17.2	52	9	09AZ71	09az71 lactococcus	188	5	17.2	143	17	0979H4	0979H4 thermoplasm
116	5	17.2	53	16	09PEP3	09pep3 xylella fas	189	5	17.2	145	4	09BS52	09BS52 homo sapien
117	5	17.2	54	16	09PDE7	09pde7 xylella fas	190	5	17.2	145	5	09GUY2	09GUY2 polyandroca
118	5	17.2	57	9	09AZ26	09az26 bacterioph	191	5	17.2	145	5	09GUY1	09GUY1 polyandroca
119	5	17.2	58	5	09GN76	09gn76 plasmodiu	192	5	17.2	145	5	094679	094679 polyandroca
120	5	17.2	60	12	068963	068963 hepatitis c	193	5	17.2	145	6	097914	097914 ovis aries
121	5	17.2	64	5	09SM55	09sm55 haemophilu	194	5	17.2	146	2	09R916	09R916 pseudomonas
122	5	17.2	65	10	09M6G5	09m6g5 medicago tr	195	5	17.2	147	12	09WMN0	09WMN0 human papil
123	5	17.2	68	4	075941	075941 homo sapien	196	5	17.2	148	17	09V252	09V252 pyrococcus
124	5	17.2	69	5	09VFW2	09vfw2 drosophila	197	5	17.2	148	17	09HJX9	09HJX9 thermoplasm
125	5	17.2	72	5	09UAB0	09uab0 conus livid	198	5	17.2	149	2	09AMJ2	09AMJ2 pantoea agg
126	5	17.2	77	3	006740	006740 saccharomy	199	5	17.2	151	5	09N5B4	09N5B4 caenorhabdi
127	5	17.2	78	2	09RBE3	09rbe3 arthrobacte	200	5	17.2	151	16	098J90	098J90 rhizobium l
128	5	17.2	79	16	098A33	098a33 rhizobium l	201	5	17.2	151	17	098J90	098J90 aeropyrum p
129	5	17.2	83	10	022172	022172 arabidopsis	202	5	17.2	152	5	09VEI6	09VEI6 drosophila
130	5	17.2	83	10	094K79	094k79 arabidopsis	203	5	17.2	152	11	0924X1	0924X1 mus musculu
131	5	17.2	83	16	09A3J8	09a3j8 caulobacter	204	5	17.2	152	16	098FJ6	098FJ6 rhizobium l
132	5	17.2	84	5	09VS20	09vs20 drosophila	205	5	17.2	152	16	098FJ6	098FJ6 streptococ
133	5	17.2	85	2	09ZHG5	09zng5 lactobacill	206	5	17.2	152	17	09HTI9	09HTI9 thermoplasm
134	5	17.2	88	16	09KD32	09kd32 bacillus ha	207	5	17.2	153	6	095M63	095M63 macaca radi
135	5	17.2	89	6	029370	029370 sus scrofa	208	5	17.2	153	16	097FW1	097FW1 clostridium
136	5	17.2	93	12	069373	069373 cercopithe	209	5	17.2	154	3	09C0S4	09C0S4 cryptococcu
137	5	17.2	94	9	080105	080105 bacterioph	210	5	17.2	156	5	09BL58	09BL58 caenorhabdi
138	5	17.2	94	9	021871	021871 bacterioph	211	5	17.2	158	10	09C505	09C505 arabidopsis
139	5	17.2	94	9	037858	037858 bacterioph	212	5	17.2	158	10	094JU4	094JU4 arabidopsis
140	5	17.2	99	12	091FX7	091fx7 chilo iride	213	5	17.2	158	16	0926X6	0926X6 chlamydia p
141	5	17.2	99	16	091044	091044 pseudomonas	214	5	17.2	159	2	09S1V8	09S1V8 streptomyc
142	5	17.2	101	17	09AJC9	09ajc9 pseudomonas	215	5	17.2	159	6	09MZL0	09MZL0 canis famli
143	5	17.2	103	17	096XL3	096xl3 sulfolobus	216	5	17.2	159	10	09SC12	09SC12 senecio ver
144	5	17.2	110	16	006320	006320 mycobacteri	217	5	17.2	159	10	094SF4	094SF4 medicago sa
145	5	17.2	112	2	P77235	P77235 escherichia	218	5	17.2	159	10	093VP3	093VP3 arabidopsis
146	5	17.2	114	5	095WS4	095ws4 haemophilu	219	5	17.2	160	5	09YV19	09YV19 drosophila
147	5	17.2	116	2	069105	069105 bacillus sp	220	5	17.2	160	5	P91046	P91046 caenorhabdi
148	5	17.2	117	5	09TYE2	09tye2 polyandroca	221	5	17.2	161	2	092400	092400 pseudomonas
149	5	17.2	117	16	09A4T9	09a4t9 caulobacter	222	5	17.2	161	4	09UE66	09UE66 homo sapien
150	5	17.2	118	6	09XSS0	09xss0 canis famli	223	5	17.2	161	16	09A8T4	09A8T4 caulobacter
151	5	17.2	118	11	09DCH7	09dch7 mus musculu	224	5	17.2	162	6	095J12	095J12 pan troglod
152	5	17.2	119	12	09OP83	09op83 hepatitis e	225	5	17.2	164	13	09DEB3	09DEB3 fugu rubrip
153	5	17.2	119	12	09OP82	09op82 hepatitis e	226	5	17.2	164	17	09YFG6	09YFG6 aeropyrum p
154	5	17.2	119	12	09OP81	09op81 hepatitis e	227	5	17.2	168	10	096433	096433 glycine max
155	5	17.2	119	12	09OP80	09op80 hepatitis e	228	5	17.2	168	11	09P9I2	09P9I2 mus musculu
156	5	17.2	119	12	09OP79	09op79 hepatitis e	229	5	17.2	169	10	096432	096432 glycine max
157	5	17.2	119	12	09QP78	09qp78 hepatitis e	230	5	17.2	172	8	09MVU5	09MVU5 chidulum mon
158	5	17.2	120	2	069099	069099 bacillus st	231	5	17.2	174	3	094184	094184 aspergillus
159	5	17.2	121	5	022934	022934 caenorhabdi	232	5	17.2	174	16	067188	067188 aquifex aeo
160	5	17.2	122	11	09UDN5	09udn5 mus musculu	233	5	17.2	176	16	084229	084229 chlamydia t
161	5	17.2	123	4	09UC23	09uc23 homo sapien	234	5	17.2	177	5	095XU5	095XU5 caenorhabdi
162	5	17.2	123	5	001958	001958 bombyx mori	235	5	17.2	178	5	09NLM5	09NLM5 leishmania

236	5	17.2	178	10	09XGS9	09xgs9	309	5	17.2	232	4	09BVC5	09bvc5	homo sapien
237	5	17.2	179	10	09XFI8	09xfl8	310	5	17.2	232	16	09A166	09a166	streptococc
238	5	17.2	179	12	055706	055706	311	5	17.2	233	5	096213	096213	schistosoma
239	5	17.2	180	11	09CTF8	09ctf8	312	5	17.2	234	16	0988Y0	0988Y0	rhizobium 1
240	5	17.2	180	16	0915V4	0915v4	313	5	17.2	235	2	030712	030712	synechococc
241	5	17.2	183	3	09C402	09c402	314	5	17.2	235	16	053324	053324	mycobacteri
242	5	17.2	183	11	091Z82	091z82	315	5	17.2	236	9	09MBT4	09mbt4	staphylococ
243	5	17.2	185	10	09MBD6	09mbd6	316	5	17.2	237	5	09VFC6	09vfc6	dirosophila
244	5	17.2	185	10	09SFE90	09sfe90	317	5	17.2	237	8	094990	094990	myrmecia gu
245	5	17.2	185	11	099P50	099p50	318	5	17.2	237	16	0911Y3	0911Y3	pseudomonas
246	5	17.2	185	12	09JGP4	09jgp4	319	5	17.2	237	16	09AC56	09ac56	caulobacter
247	5	17.2	185	17	09Y9U1	09y9u1	320	5	17.2	238	8	09MP35	09mp35	anonymoxr
248	5	17.2	186	2	093T20	093t20	321	5	17.2	238	8	09MP29	09mp29	iridomyrmex
249	5	17.2	186	2	093T19	093t19	322	5	17.2	238	8	09MP26	09mp26	ochetellus
250	5	17.2	188	4	096AV3	096av3	323	5	17.2	239	8	09MP25	09mp25	tapinoma sp
251	5	17.2	189	2	09RJ82	09rj82	324	5	17.2	240	2	093KM1	093km1	streptomyc
252	5	17.2	190	13	091649	091649	325	5	17.2	240	10	022730	022730	arabidopsis
253	5	17.2	190	16	098B09	098b09	326	5	17.2	241	12	084617	084617	paramectum
254	5	17.2	191	5	09N9H7	09n9h7	327	5	17.2	241	16	084728	084728	chlamydia t
255	5	17.2	195	2	046086	046086	328	5	17.2	241	16	P72757	P72757	synechocyst
256	5	17.2	195	12	09J839	09j839	329	5	17.2	243	16	0921N3	0921n3	ricketsia
257	5	17.2	196	10	09SXX1	09sxx1	330	5	17.2	245	2	09LB15	09lb15	vibrio para
258	5	17.2	196	10	09AVX0	09avx0	331	5	17.2	245	8	09MF49	09mf49	beta vulgar
259	5	17.2	199	4	09H3L5	09h3l5	332	5	17.2	246	11	09Z2U8	09z2u8	mus musc
260	5	17.2	200	10	091ME5	091me5	333	5	17.2	247	2	086486	086486	streptomyc
261	5	17.2	201	2	08S165	08s165	334	5	17.2	247	16	092AK1	092ak1	listeria in
262	5	17.2	202	6	09TUF5	09tuf5	335	5	17.2	248	12	040634	040634	salimarine
263	5	17.2	202	6	09TUF4	09tuf4	336	5	17.2	249	13	0908V6	0908v6	epitretus
264	5	17.2	204	2	09MB92	09mb92	337	5	17.2	250	2	09X474	09x474	enterococc
265	5	17.2	204	2	093NM2	093nm2	338	5	17.2	250	2	09L8R8	09l8r8	bacillus sp
266	5	17.2	204	10	040023	040023	339	5	17.2	250	17	09HPD3	09hpd3	halobacteri
267	5	17.2	204	16	098L32	098l32	340	5	17.2	251	11	09D6M1	09d6m1	mus musc
268	5	17.2	206	16	09KLN1	09kln1	341	5	17.2	253	4	09BTAY	09btay	homo sapien
269	5	17.2	207	15	091014	091014	342	5	17.2	253	16	09A9J4	09a9j4	caulobacter
270	5	17.2	208	4	09H8B6	09h8b6	343	5	17.2	254	4	09BX17	09bx17	homo sapien
271	5	17.2	208	10	004527	004527	344	5	17.2	254	10	094EV4	094ev4	zea mays (m
272	5	17.2	208	17	097BB4	097bb4	345	5	17.2	254	16	09KMU4	09kmu4	vibrio chol
273	5	17.2	209	2	093CX7	093cx7	346	5	17.2	255	2	083KZ6	083kz6	burkholderi
274	5	17.2	209	5	09V5Z7	09v5z7	347	5	17.2	256	16	08KSU9	08ksu9	vibrio chol
275	5	17.2	210	5	09XT12	09xt12	348	5	17.2	256	16	09KRF7	09krf7	vibrio chol
276	5	17.2	210	8	037662	037662	349	5	17.2	257	2	051720	051720	propionibac
277	5	17.2	211	10	09LUR8	09lur8	350	5	17.2	257	16	098FG3	098fg3	rhizobium 1
278	5	17.2	212	2	099Q39	099q39	351	5	17.2	258	4	09NSY0	09nsy0	homo sapien
279	5	17.2	213	2	09YPM9	09ypm9	352	5	17.2	258	11	091V36	091v36	mus musc
280	5	17.2	213	16	092NB2	092nb2	353	5	17.2	258	16	09RYV6	09ryv6	deinococcus
281	5	17.2	214	6	09BE68	09be68	354	5	17.2	259	4	09MTQ1	09mtq1	homo sapien
282	5	17.2	214	11	P97712	P97712	355	5	17.2	259	15	09NBA6	09nba6	glossina mo
283	5	17.2	215	5	096317	096317	356	5	17.2	259	16	076583	076583	human immun
284	5	17.2	215	16	098DX2	098dx2	357	5	17.2	259	15	050181	050181	pseudomonas
285	5	17.2	216	10	09LQY9	09lqy9	358	5	17.2	260	4	09H3N5	09h3n5	homo sapien
286	5	17.2	217	5	09UOV5	09uov5	359	5	17.2	260	4	09S162	09s162	homo sapien
287	5	17.2	218	13	013162	013162	360	5	17.2	260	5	09W1H6	09w1h6	dirosophila
288	5	17.2	218	16	09KYV8	09kyv8	361	5	17.2	261	5	09W3C6	09w3c6	dirosophila
289	5	17.2	220	2	09RJ11	09rj11	362	5	17.2	261	15	078279	078279	human immun
290	5	17.2	220	2	066045	066045	363	5	17.2	262	10	091K10	091k10	populus x c
291	5	17.2	221	2	09RCX5	09rcx5	364	5	17.2	263	5	09U215	09u215	caenorhabdi
292	5	17.2	221	16	09HV32	09hv32	365	5	17.2	264	2	08S946	08s946	spingomona
293	5	17.2	221	16	09Z015	09z015	366	5	17.2	264	4	09BR37	09br37	homo sapien
294	5	17.2	224	5	09WK8	09wk8	367	5	17.2	265	10	09AUS7	09aus7	oryza sativ
295	5	17.2	224	16	0987W0	0987w0	368	5	17.2	266	6	069366	069366	rhodococcus
296	5	17.2	225	4	012918	012918	369	5	17.2	266	11	09R2A7	09r2a7	mus musc
297	5	17.2	225	10	09AY88	09ay88	370	5	17.2	266	12	0919Y9	0919y9	culex nigrl
298	5	17.2	225	11	099KQ3	099kq3	371	5	17.2	266	16	098C07	098c07	rhizobium 1
299	5	17.2	225	16	097D77	097d77	372	5	17.2	267	2	0914U4	0914u4	streptomyc
300	5	17.2	226	2	09AH05	09ah05	373	5	17.2	267	10	0945M2	0945m2	arabidopsis
301	5	17.2	227	5	P91288	P91288	374	5	17.2	267	2	P70831	P70831	borrelia bu
302	5	17.2	230	2	09RI61	09ri61	375	5	17.2	267	16	09CWM2	09cwm2	pasteurella
303	5	17.2	230	16	09PMA1	09pma1	376	5	17.2	267	17	09HWM1	09hwm1	halobacteri
304	5	17.2	230	16	09HXR5	09hxr5	377	5	17.2	268	2	09AC82	09ac82	yersinia pe
305	5	17.2	231	16	09AB05	09ab05	378	5	17.2	269	4	09BRE6	09bre6	homo sapien
306	5	17.2	231	16	09BDF0	09bdf0	379	5	17.2	269	11	091YV8	091yv8	mus musc
307	5	17.2	231	16	092D04	092d04	380	5	17.2	271	6	09N2J4	09n2j4	canis fami
308	5	17.2	232	2	P71502	P71502	381	5	17.2	271	10	0942K1	0942k1	oryza sativ

382	5	17.2	271	16	092ST4	092S14 rhizobium m	455	5	17.2	302	2	0926E0	0926E0 streptococc
383	5	17.2	272	5	096UAI	096UAI trypanosoma	456	5	17.2	302	17	09V296	09V296 pyrococcus
384	5	17.2	272	5	096UAI	096UAI trypanosoma	457	5	17.2	303	5	017706	017706 caenorhabdi
385	5	17.2	273	11	09D3C8	0933C8 mus musculu	458	5	17.2	303	12	09Q1U9	09Q1U9 salmeline
386	5	17.2	273	16	006162	006162 mycobacteri	459	5	17.2	303	16	09PDL9	09PDL9 xylella fas
387	5	17.2	273	16	049993	049993 mycobacteri	460	5	17.2	304	16	09AB87	09AB87 caulobacter
388	5	17.2	273	16	092JG5	092JG5 rickettsia	461	5	17.2	305	5	09U2Y9	09U2Y9 caenorhabdi
389	5	17.2	275	16	097NFI	097NFI streptococc	462	5	17.2	307	16	09A976	09A976 caulobacter
390	5	17.2	276	4	09BX18	09BX18 homo sapien	463	5	17.2	308	2	050894	050894 myxococcus
391	5	17.2	276	5	09GU99	09GU99 trypanosoma	464	5	17.2	309	4	09BW75	09BW75 homo sapien
392	5	17.2	276	5	044029	044029 trypanosoma	465	5	17.2	309	4	096F11	096F11 homo sapien
393	5	17.2	276	11	09E034	09E034 mus musculu	466	5	17.2	309	10	0982075	0982075 achlya ambi
394	5	17.2	276	12	09J560	09J560 fowipox vir	467	5	17.2	309	10	09SAS9	09SAS9 achlya ambi
395	5	17.2	277	12	09J316	09J316 saccharomyc	468	5	17.2	310	4	09BX67	09BX67 homo sapien
397	5	17.2	277	12	09JF57	09JF57 vaccinia vi	469	5	17.2	310	10	0943S1	0943S1 oryza sativ
398	5	17.2	277	16	09KB60	09KB60 bacillus ha	470	5	17.2	310	11	09EPK4	09EPK4 mus musculu
399	5	17.2	278	2	P95551	P95551 pseudomonas	471	5	17.2	310	11	09DB87	09DB87 mus musculu
400	5	17.2	278	4	09BRZ2	09BRZ2 homo sapien	472	5	17.2	310	11	09D1M9	09D1M9 mus musculu
401	5	17.2	278	4	09NV08	09NV08 homo sapien	473	5	17.2	310	16	09CNI7	09CNI7 pasteurilla
402	5	17.2	278	13	013096	013096 salmo trutt	474	5	17.2	311	2	09R6P0	09R6P0 agrobacteri
403	5	17.2	280	12	056299	056299 human herpe	475	5	17.2	311	2	093SFS	093SFS eubacterium
404	5	17.2	280	12	069512	069512 human herpe	476	5	17.2	311	2	0943R8	0943R8 oryza sativ
405	5	17.2	280	16	069742	069742 mycobacteri	477	5	17.2	312	11	09KIB3	09KIB3 porphyromon
406	5	17.2	280	17	029933	029933 archaeoglob	478	5	17.2	312	11	09D9Z1	09D9Z1 methanosarc
407	5	17.2	281	2	09RN66	09RN66 streptomyce	479	5	17.2	313	1	09HH19	09HH19 oryza sativ
408	5	17.2	281	5	09V3A9	09V3A9 dirosophila	480	5	17.2	314	10	094LM3	094LM3 oryza sativ
409	5	17.2	283	2	085856	085856 sphingomona	481	5	17.2	314	11	09CSF8	09CSF8 mus musculu
410	5	17.2	283	2	09A1Y5	09A1Y5 candidatus	482	5	17.2	314	16	09KKQ9	09KKQ9 vibrio chol
411	5	17.2	283	6	0951M3	0951M3 macaca fasc	483	5	17.2	315	16	083312	083312 treponema p
412	5	17.2	283	16	0927Y1	0927Y1 listeria in	484	5	17.2	316	16	099YQ1	099YQ1 streptococc
413	5	17.2	284	10	094BH3	094BH3 drosiera com	485	5	17.2	317	4	09NPL6	09NPL6 homo sapien
414	5	17.2	284	12	0919P1	0919P1 culcx nigri	486	5	17.2	318	16	09KMT1	09KMT1 vibrio chol
415	5	17.2	285	11	0921H0	0921H0 mus musculu	487	5	17.2	318	17	058334	058334 pyrococcus
416	5	17.2	286	10	094BP1	094BP1 arabadidopsis	488	5	17.2	319	2	005423	005423 mycobacteri
417	5	17.2	286	10	094BP1	094BP1 cucurbita s	489	5	17.2	319	5	017914	017914 caenorhabdi
418	5	17.2	287	10	042564	042564 arabadidopsis	490	5	17.2	319	16	09RWM0	09RWM0 deinococcus
419	5	17.2	287	10	091ND6	091ND6 arabadidopsis	491	5	17.2	320	5	096U49	096U49 dirosophila
420	5	17.2	287	10	081810	081810 arabadidopsis	492	5	17.2	320	5	09N9Q4	09N9Q4 leishmania
421	5	17.2	288	10	09XG58	09XG58 zantedesch	493	5	17.2	320	5	091261	091261 caenorhabdi
422	5	17.2	288	10	093780	093780 gossypium h	494	5	17.2	320	12	09JGV2	09JGV2 plutella xy
423	5	17.2	289	4	096RJ7	096RJ7 homo sapien	495	5	17.2	320	17	09UZU3	09UZU3 pyrococcus
424	5	17.2	289	4	096MP8	096MP8 homo sapien	496	5	17.2	322	10	094BM9	094BM9 arabadidopsis
425	5	17.2	289	10	09S717	09S717 arabadidopsis	497	5	17.2	322	16	091310	091310 borella bu
426	5	17.2	289	10	09ZT26	09ZT26 zea mays (m	498	5	17.2	324	2	09S5E9	09S5E9 synchococc
427	5	17.2	289	16	09CER0	09CER0 lactococcus	499	5	17.2	324	2	09AQ17	09AQ17 bradyrhizob
428	5	17.2	290	2	057103	057103 yersinia ps	500	5	17.2	324	2	09RJ00	09RJ00 streptomyce
429	5	17.2	290	4	09X3Z8	09X3Z8 homo sapien	501	5	17.2	324	10	093163	093163 arabadidopsis
430	5	17.2	290	8	096157	096157 humularia c	502	5	17.2	324	10	09C8M9	09C8M9 arabadidopsis
431	5	17.2	290	10	023139	023139 arabadidopsis	503	5	17.2	325	3	09HEX3	09HEX3 cochlilobu
432	5	17.2	290	12	091K76	091K76 spodoptera	504	5	17.2	326	5	09XW6	09XW6 caenorhabdi
433	5	17.2	291	10	0941C3	0941C3 hordeum vul	505	5	17.2	326	10	094JG3	094JG3 oryza sativ
434	5	17.2	291	17	09V189	09V189 pyrococcus	506	5	17.2	327	8	09G164	09G164 paramachea
435	5	17.2	292	16	09CEG5	09CEG5 lactococcus	507	5	17.2	327	11	099JK6	099JK6 mus musculu
436	5	17.2	295	16	09CPA0	09CPA0 pasteurilla	508	5	17.2	327	16	098UJ6	098UJ6 rhizobium l
437	5	17.2	296	10	09ZUH5	09ZUH5 arabadidopsis	509	5	17.2	328	2	086149	086149 campylobact
438	5	17.2	296	10	09M2E2	09M2E2 arabadidopsis	510	5	17.2	328	12	068387	068387 human cytom
439	5	17.2	296	16	09KT82	09KT82 vibrio chol	511	5	17.2	328	16	09PN63	09PN63 campylobact
440	5	17.2	297	5	020803	020803 caenorhabdi	512	5	17.2	329	2	09RL37	09RL37 streptomyce
441	5	17.2	297	12	P87725	P87725 garlic late	513	5	17.2	329	16	0929Z7	0929Z7 listeria in
442	5	17.2	298	2	093RY0	093RY0 streptomyce	514	5	17.2	330	2	09K5D5	09K5D5 campylobact
443	5	17.2	299	10	091ND9	091ND9 arabadidopsis	515	5	17.2	330	4	09NT67	09NT67 schizosach
444	5	17.2	299	10	091ND8	091ND8 arabadidopsis	516	5	17.2	330	4	09FL50	09FL50 homo sapien
445	5	17.2	299	11	09FPD5	09FPD5 arabadidopsis	517	5	17.2	330	10	09FL50	09FL50 arabadidopsis
446	5	17.2	299	11	09Q0U1	09Q0U1 mus musculu	518	5	17.2	331	16	066718	066718 aquilex aeo
447	5	17.2	299	11	09CZC1	09CZC1 mus musculu	519	5	17.2	332	5	095W18	095W18 caenorhabdi
448	5	17.2	299	11	09OYU1	09OYU1 rattus ratt	520	5	17.2	332	8	09TJZ8	09TJZ8 plactetia ac
449	5	17.2	300	4	09H704	09H704 homo sapien	521	5	17.2	332	16	0916C6	0916C6 pseudomonas
450	5	17.2	300	11	035980	035980 mus musculu	522	5	17.2	333	8	020001	020001 aciphylla a
451	5	17.2	301	4	096EH1	096EH1 homo sapien	523	5	17.2	333	8	034833	034833 kluyveromyc
452	5	17.2	301	6	09XJG0	09XJG0 bacterioph	524	5	17.2	333	8	095FA3	095FA3 myodocarpus
453	5	17.2	301	9	09PFU2	09PFU2 xylella fas	525	5	17.2	335	16	09KT36	09KT36 vibrio chol
454	5	17.2	301	16	092TW9	092TW9 rhizobium m	526	5	17.2	335	16	09PM7	09PM7 campylobact
							527	5	17.2	335	16	0916B9	0916B9 pseudomonas

528	5	17.2	336	13	Q9DDQ8	Q9ddg8 xenopus lae	601	5	17.2	358	8	Q95F99	Q95f99 mydocarpus
529	5	17.2	336	16	Q9KSQ2	Q9ksq2 vibrio chol	602	5	17.2	358	8	Q95F98	Q95f98 mydocarpus
530	5	17.2	337	2	Q9XDN2	Q9xdn2 salmonella	603	5	17.2	358	16	Q9KMN8	Q9kmt8 vibrio chol
531	5	17.2	337	4	Q96NKR	Q96nkr homo sapien	604	5	17.2	359	5	Q9V6C2	Q9vec2 drosophila
532	5	17.2	337	10	Q9ZT53	Q9zt53 griffithsia	605	5	17.2	359	8	Q20010	Q20010 anginon rug
533	5	17.2	338	2	Q30542	Q30542 agrobacteri	606	5	17.2	359	8	Q20020	Q20020 carum carvi
534	5	17.2	338	16	Q9PPT2	Q9pdt2 xylella fas	607	5	17.2	359	8	Q20054	Q20054 heteromorph
535	5	17.2	338	16	Q9AAU4	Q9aa4 caulobacter	608	5	17.2	359	8	Q20062	Q20062 melanophyll
536	5	17.2	339	2	Q9R6D8	Q9r6d8 agrobacteri	609	5	17.2	359	8	Q20005	Q20005 ammi majus
537	5	17.2	339	8	Q9TJ29	Q9tj29 ormocarpum	610	5	17.2	359	8	Q95F90	Q95f90 torellia se
538	5	17.2	339	12	Q90148	Q90148 rachiplusia	611	5	17.2	359	12	Q67620	Q67620 tomato yell
539	5	17.2	339	16	Q9KSB3	Q9ksb3 vibrio chol	612	5	17.2	359	13	Q9DE48	Q9de48 brachydanio
540	5	17.2	340	2	Q49592	Q49592 mycobacteri	613	5	17.2	360	2	Q68551	Q68551 pseudomonas
541	5	17.2	341	5	Q01466	Q01466 caenorhabdi	614	5	17.2	360	8	Q20041	Q20041 eryngium bo
542	5	17.2	341	5	Q0R9Y4	Q9r9y4 pseudomonas	615	5	17.2	360	8	Q20086	Q20086 sanicula gr
543	5	17.2	341	2	Q85214	Q85214 pseudomonas	616	5	17.2	360	8	Q20043	Q20043 eremocharis
544	5	17.2	341	3	Q74509	Q74509 schizosacch	617	5	17.2	360	8	Q20057	Q20057 kalopanax p
545	5	17.2	341	5	Q9VR72	Q9vr72 drosophila	618	5	17.2	360	8	Q20065	Q20065 munroidendr
546	5	17.2	342	2	Q87333	Q87333 campylobact	619	5	17.2	360	8	Q20093	Q20093 tetraplasan
547	5	17.2	342	5	Q9UBV7	Q9ubv7 brachyosteo	620	5	17.2	360	8	Q20013	Q20013 acanthopana
548	5	17.2	342	10	Q9FTQ7	Q9ftq7 oryza sativ	621	5	17.2	360	8	Q20008	Q20008 arthropophyll
549	5	17.2	342	16	Q67713	Q67713 aquifex aeo	622	5	17.2	360	8	Q20036	Q20036 cheirodendr
550	5	17.2	342	16	Q83923	Q83923 treponema p	623	5	17.2	360	8	Q20037	Q20037 dendropanax
551	5	17.2	343	5	Q9XV25	Q9xv25 caenorhabdi	624	5	17.2	360	8	Q20045	Q20045 fatsia japo
552	5	17.2	343	17	Q97228	Q97228 sulfolobus	625	5	17.2	360	8	Q20052	Q20052 hederia heli
553	5	17.2	344	5	Q61754	Q61754 caenorhabdi	626	5	17.2	360	8	Q20067	Q20067 meryta sinc
554	5	17.2	344	10	Q48727	Q48727 arabidopsis	627	5	17.2	360	8	Q20074	Q20074 polyscias g
555	5	17.2	344	10	Q9ARW2	Q9arw2 oryza sativ	628	5	17.2	360	8	Q20071	Q20071 pseudopanax
556	5	17.2	345	2	Q32708	Q32708 kluyvera ci	629	5	17.2	360	8	Q20082	Q20082 reynoldsia
557	5	17.2	345	5	Q20378	Q20378 caenorhabdi	630	5	17.2	360	8	Q20083	Q20083 schefflera
558	5	17.2	345	8	Q9GW2	Q9gw2 anabeta dome	631	5	17.2	360	8	Q20095	Q20095 trevesia pa
559	5	17.2	346	2	P70801	P70801 anabaena va	632	5	17.2	360	8	Q20092	Q20092 tupidanthus
560	5	17.2	346	5	Q17086	Q17086 caenorhabdi	633	5	17.2	360	8	Q20053	Q20053 heracleum l
561	5	17.2	346	16	Q981F1	Q981f1 rhizobium l	634	5	17.2	360	8	Q95F96	Q95f96 astrotricha
562	5	17.2	347	2	Q9RL31	Q9rl31 streptomyce	635	5	17.2	360	8	Q95F95	Q95f95 astrotricha
563	5	17.2	347	12	Q91FD6	Q91fd6 chilo iride	636	5	17.2	360	8	Q95F91	Q95f91 arctopus ec
564	5	17.2	348	10	Q9LEW5	Q9lew5 arabidopsis	637	5	17.2	360	8	Q95F89	Q95f89 dichoscladi
565	5	17.2	349	2	Q9ZF42	Q9zf42 lactococcus	638	5	17.2	360	8	Q95F88	Q95f88 diplaspis h
566	5	17.2	349	10	Q9FLV7	Q9flv7 arabidopsis	639	5	17.2	361	5	Q9N5E7	Q9n5e7 caenorhabdi
567	5	17.2	349	16	Q9CFG5	Q9cfg5 lactococcus	640	5	17.2	361	8	Q20003	Q20003 anethum gra
568	5	17.2	350	16	Q9RW72	Q9rt72 delnoccocus	641	5	17.2	361	8	Q20004	Q20004 angelica lu
569	5	17.2	351	2	Q44630	Q44630 bruceia ab	642	5	17.2	361	8	Q20009	Q20009 arracacia q
570	5	17.2	351	3	Q96VDA	Q96vda ustilago ma	643	5	17.2	361	8	Q20015	Q20015 bifora amer
571	5	17.2	351	4	Q9H4N9	Q9h4n9 homo sapien	644	5	17.2	361	8	Q20021	Q20021 cicuta doug
572	5	17.2	351	8	Q20069	Q20069 oreonana pu	645	5	17.2	361	8	Q20026	Q20026 coxana pur
573	5	17.2	351	10	Q9FX40	Q9fx40 arabidopsis	646	5	17.2	361	8	Q20033	Q20033 coriandrum
574	5	17.2	351	10	Q94GG9	Q94gg9 oryza sativ	647	5	17.2	361	8	Q20039	Q20039 daucus caro
575	5	17.2	351	12	Q82676	Q82676 indian cass	648	5	17.2	361	8	Q20038	Q20038 donnellsmi
576	5	17.2	352	16	Q69627	Q69627 mycobacteri	649	5	17.2	361	8	Q20044	Q20044 enantiophyl
577	5	17.2	353	2	Q93CP8	Q93cp8 shigella fl	650	5	17.2	361	8	Q20042	Q20042 endressia c
578	5	17.2	353	8	Q33791	Q33791 alligator m	651	5	17.2	361	8	Q20046	Q20046 foeniculum
579	5	17.2	354	8	Q95F94	Q95f94 osmoxylon g	652	5	17.2	361	8	Q20061	Q20061 levisticum
580	5	17.2	354	8	Q95F93	Q95f93 osmoxylon n	653	5	17.2	361	8	Q20059	Q20059 lomatum ho
581	5	17.2	355	2	Q9LLA1	Q9lla1 streptomyce	654	5	17.2	361	8	Q20072	Q20072 petroselinu
582	5	17.2	355	8	Q20091	Q20091 tauschia ar	655	5	17.2	361	8	Q20081	Q20081 rhodosciadi
583	5	17.2	355	8	Q95FA6	Q95fa6 mydocarpus	656	5	17.2	361	8	Q20089	Q20089 scandix pec
584	5	17.2	356	8	Q20000	Q20000 araccacia a	657	5	17.2	361	8	Q20085	Q20085 spermolep
585	5	17.2	357	4	Q96Q09	Q96q09 homo sapien	658	5	17.2	361	8	Q20084	Q20084 sphegnosclad
586	5	17.2	357	8	Q20049	Q20049 gastonia ro	659	5	17.2	361	8	Q20094	Q20094 taenidia in
587	5	17.2	357	13	Q9DE05	Q9de05 raja erinac	660	5	17.2	362	3	Q02417	Q02417 saccharomyc
588	5	17.2	358	5	Q9U220	Q9u220 caenorhabdi	661	5	17.2	362	3	Q03953	Q03953 saccharomyc
589	5	17.2	358	8	Q20022	Q20022 centella er	662	5	17.2	362	3	Q96TS6	Q96ts6 phlebia rad
590	5	17.2	358	8	Q20012	Q20012 aralia spin	663	5	17.2	362	8	Q95F92	Q95f92 actinotus h
591	5	17.2	358	8	Q20063	Q20063 macknlaya	664	5	17.2	362	12	Q87020	Q87020 killer viru
592	5	17.2	358	8	Q20075	Q20075 panax quinq	665	5	17.2	362	16	Q9CN26	Q9cn26 pasteurilla
593	5	17.2	358	8	Q20051	Q20051 hymenosporu	666	5	17.2	363	8	Q20060	Q20060 laserpitium
594	5	17.2	358	8	Q20079	Q20079 pittosporum	667	5	17.2	363	8	Q20077	Q20077 pimpliella
595	5	17.2	358	8	Q20087	Q20087 sollya hete	668	5	17.2	363	8	Q20006	Q20006 aralidium p
596	5	17.2	358	8	Q95FB1	Q95fb1 macknlaya	669	5	17.2	363	8	Q9MSE0	Q9mse0 drosera cap
597	5	17.2	358	8	Q95FA5	Q95fa5 mydocarpus	670	5	17.2	364	8	Q20024	Q20024 ciclospermu
598	5	17.2	358	8	Q95FA4	Q95fa4 mydocarpus	671	5	17.2	364	8	Q20068	Q20068 neogoezia m
599	5	17.2	358	8	Q95FA2	Q95fa2 mydocarpus	672	5	17.2	364	8	Q20070	Q20070 oenanthe sa
600	5	17.2	358	8	Q95FA1	Q95fa1 mydocarpus	673	5	17.2	364	8	Q20073	Q20073 perideridia

674	5	17.2	364	8	O20088	O20088 shoshonea p	747	5	17.2	397	2	Q9X3W2	Q9x3w2 zymomonas m
675	5	17.2	364	8	O20090	O20090 sium suave.	748	5	17.2	397	5	Q9NKF2	Q9nxf2 drosophila
676	5	17.2	364	11	O9CTP6	O9ctp6 mus musculus	749	5	17.2	397	6	Q9XSX7	Q9xsx7 bos taurus
677	5	17.2	365	3	Q9P3Z5	Q9p3z5 coriolus ve	750	5	17.2	397	16	Q92BD4	Q92bd4 listeria in
678	5	17.2	365	3	Q12720	Q12720 tranetes ve	751	5	17.2	398	3	O13358	O13358 aspergillus
679	5	17.2	365	3	Q99058	Q99058 tranetes ve	752	5	17.2	398	4	Q9NUJ0	Q9nuj0 homo sapien
680	5	17.2	365	16	P73132	P73132 synchocyst	753	5	17.2	398	17	Q9HKN3	Q9hkn3 thermoplasm
681	5	17.2	367	8	Q9T390	Q9t390 nephroselmi	754	5	17.2	400	5	Q9GZH7	Q9gzh7 caenorhabdi
682	5	17.2	367	8	Q95PB2	Q95tb2 apiopetalum	755	5	17.2	400	16	Q99TF2	Q99tf2 staphylococ
683	5	17.2	369	2	Q9AD01	Q9ad01 streptomyce	756	5	17.2	400	16	Q931P6	Q931p6 staphylococ
684	5	17.2	369	5	Q9N2T7	Q9n2t7 caenorhabdi	757	5	17.2	401	10	Q9FS87	Q9fs87 solanum tub
685	5	17.2	369	10	Q9FJ25	Q9fj25 arabidopsis	758	5	17.2	401	11	Q9CWC8	Q9cwc8 mus musculu
686	5	17.2	369	12	Q9QVJ1	Q9qvj1 borna disea	759	5	17.2	401	11	Q91ZG3	Q91zg3 mus musculu
687	5	17.2	369	16	Q9KPU3	Q9kpu3 vibrio chol	760	5	17.2	402	2	Q9F8U7	Q9f8u7 streptomyce
688	5	17.2	370	5	Q9VA44	Q9va44 drosophila	761	5	17.2	403	12	Q9YV17	Q9yvi7 melanoplus
689	5	17.2	370	12	O10392	O10392 borna disea	762	5	17.2	405	16	Q9WYN1	Q9wyn1 thermotoga
690	5	17.2	370	12	O10398	O10398 borna disea	763	5	17.2	405	16	Q9KQJ1	Q9kqj1 vibrio chol
691	5	17.2	370	12	O10401	O10401 borna disea	764	5	17.2	408	2	Q9RN55	Q9rn55 streptomyce
692	5	17.2	370	12	Q91UL8	Q91ul8 borna disea	765	5	17.2	408	10	Q9SM62	Q9sm62 pisum sativ
693	5	17.2	370	12	Q91UL4	Q91ul4 borna disea	766	5	17.2	408	10	Q9ASA9	Q9asa9 oryza sativ
694	5	17.2	370	12	Q910H1	Q910h1 borna disea	767	5	17.2	409	10	Q9SM61	Q9sm61 pisum sativ
695	5	17.2	370	16	Q9A3Q8	Q9a3q8 caulobacter	768	5	17.2	409	10	Q9SWG0	Q9swg0 arabidopsis
696	5	17.2	370	16	Q92XLI	Q92xli rhizobium m	769	5	17.2	409	16	Q914C6	Q914c6 pseudomonas
697	5	17.2	371	2	Q9LA72	Q9la72 thiolobacillu	770	5	17.2	411	16	Q9Z8V5	Q9z8v5 chlamydia p
698	5	17.2	371	17	Q9HSF1	Q9hsf1 halobacteri	771	5	17.2	411	16	Q9KR84	Q9kr84 vibrio chol
699	5	17.2	372	5	Q9U1I8	Q9u1i8 drosophila	772	5	17.2	411	16	Q9CCM2	Q9ccm2 mycobacteri
700	5	17.2	372	5	O17208	O17208 caenorhabdi	773	5	17.2	411	16	Q9A3L1	Q9a3l1 caulobacter
701	5	17.2	372	5	Q95W17	Q95w17 caenorhabdi	774	5	17.2	412	10	Q9FS88	Q9fs88 solanum tub
702	5	17.2	373	3	Q94661	Q94661 schizosacch	775	5	17.2	413	2	Q9S0S5	Q9s0s5 streptococc
703	5	17.2	374	16	Q9PKY2	Q9pky2 chlamydia m	776	5	17.2	413	2	Q9ALW4	Q9alw4 streptococc
704	5	17.2	374	16	Q99059	Q99u59 staphylococ	777	5	17.2	413	2	Q9AFG9	Q9afg9 streptococc
705	5	17.2	374	16	Q931R7	Q931r7 staphylococ	778	5	17.2	413	2	Q93TI0	Q93ti0 streptococc
706	5	17.2	374	16	Q92AN4	Q92an4 listeria in	779	5	17.2	415	4	O43266	O43266 homo sapien
707	5	17.2	376	16	Q9CMR3	Q9cmr3 pasteurella	780	5	17.2	416	4	O43153	O43153 homo sapien
708	5	17.2	376	16	Q92UG0	Q92ug0 rhizobium m	781	5	17.2	416	5	Q18945	Q18945 caenorhabdi
709	5	17.2	377	5	Q9GPI3	Q9gpi3 ixodes ric	782	5	17.2	416	16	Q55428	Q55428 synchocyst
710	5	17.2	377	12	Q99GV8	Q99gv8 helicoverp	783	5	17.2	416	16	Q9KET7	Q9ket7 bacillus ha
711	5	17.2	377	12	Q91BV6	Q91bv6 helicoverpa	784	5	17.2	419	8	O20202	O20202 panax ginse
712	5	17.2	377	16	Q929E9	Q929e9 listeria in	785	5	17.2	419	8	O20203	O20203 panax japon
713	5	17.2	378	2	Q9F7B3	Q9f7b3 salmonella	786	5	17.2	419	8	O20204	O20204 panax quinq
714	5	17.2	378	8	Q9NSE2	Q9mse2 drosera reg	787	5	17.2	419	8	Q9MT79	Q9mt79 panax vietn
715	5	17.2	378	8	Q9GDE5	Q9gdz5 aesclynomen	788	5	17.2	419	10	Q9ES34	Q9fs34 panax notog
716	5	17.2	379	5	O02300	O02300 caenorhabdi	789	5	17.2	420	10	Q9FS33	Q9fs33 streptomyce
717	5	17.2	379	16	Q9KF62	Q9kf62 bacillus ha	790	5	17.2	420	2	Q9FX27	Q9fx27 streptomyce
718	5	17.2	379	16	Q9HX58	Q9hx58 pseudomonas	791	5	17.2	420	3	Q9PS53	Q9ps53 neuropeptora
719	5	17.2	380	1	Q50515	Q50515 methanother	792	5	17.2	420	5	Q9VSL9	Q9vsl9 drosophila
720	5	17.2	382	11	Q9JKU8	Q9jku8 mus musculu	793	5	17.2	420	11	Q99JT10	Q99j10 mus musculu
721	5	17.2	383	10	Q9SUR8	Q9sur8 arabidopsis	794	5	17.2	421	3	O43068	O43068 schizosacch
722	5	17.2	383	10	Q9SS84	Q9ss84 arabidopsis	795	5	17.2	421	10	P92937	P92937 arabidopsis
723	5	17.2	383	17	O27210	O27210 methanother	796	5	17.2	421	10	Q9LZW5	Q9lzw5 arabidopsis
724	5	17.2	383	17	O57967	O57967 pyrococcus	797	5	17.2	422	5	Q9VLZ2	Q9vlz2 drosophila
725	5	17.2	383	17	Q9V266	Q9v266 pyrococcus	798	5	17.2	423	4	Q96AF6	Q96af6 homo sapien
726	5	17.2	384	5	O44753	O44753 caenorhabdi	799	5	17.2	423	5	O61532	O61532 drosophila
727	5	17.2	384	10	Q9SR98	Q9sr98 arabidopsis	800	5	17.2	423	16	Q9CIU8	Q9ciu8 lactococcus
728	5	17.2	386	12	Q9RW77	Q9rw77 melanoplus	801	5	17.2	424	2	Q9EUJ0	Q9eujo salmonella
729	5	17.2	387	10	Q93833	Q93833 arabidopsis	802	5	17.2	424	2	Q93QG6	Q93qg6 salmonella
730	5	17.2	387	17	Q9HMR5	Q9hmr5 halobacteri	803	5	17.2	424	4	Q96LX8	Q96lx8 homo sapien
731	5	17.2	388	2	O82893	O82893 escherichia	804	5	17.2	424	5	O44656	O44656 caenorhabdi
732	5	17.2	388	5	O44330	O44330 manduca sex	805	5	17.2	424	12	Q91HZ2	Q91hz2 bovine vita
733	5	17.2	389	4	Q9UDV9	Q9udv9 homo sapien	806	5	17.2	425	5	Q9WZG4	Q9wz4 drosophila
734	5	17.2	390	4	O43842	O43842 homo sapien	807	5	17.2	426	5	Q9NE22	Q9nfe2 caenorhabdi
735	5	17.2	390	4	Q9BU90	Q9bu90 homo sapien	808	5	17.2	426	16	Q9KUG3	Q9kug3 vibrio chol
736	5	17.2	390	16	Q98CB9	Q98cb9 rhizobium l	809	5	17.2	427	5	O77101	O77101 aedes aegyp
737	5	17.2	390	16	Q92QY8	Q92qy8 rhizobium m	810	5	17.2	428	16	O917A3	O917a3 pseudomonas
738	5	17.2	391	2	Q93CM0	Q93cm0 enterococcu	811	5	17.2	428	16	O9A3U9	O9a3u9 caulobacter
739	5	17.2	394	5	Q9VKJ7	Q9vkj7 drosophila	812	5	17.2	428	16	Q92T40	Q92t40 rhizobium m
740	5	17.2	394	16	Q9RSC3	Q9rsc3 delnococcus	813	5	17.2	429	5	Q91V77	Q91v77 drosophila
741	5	17.2	395	2	Q93MK5	Q93mk5 vibrio vuln	814	5	17.2	430	2	Q9REU1	Q9reu1 streptomyce
742	5	17.2	395	8	Q9TK00	Q9tk00 ormocarpum	815	5	17.2	430	2	Q9JN56	Q9jn56 streptomyce
743	5	17.2	395	16	Q983J8	Q983j8 rhizobium l	816	5	17.2	430	4	Q9H1Q0	Q9h1q0 homo sapien
744	5	17.2	396	5	O27463	O27463 caenorhabdi	817	5	17.2	430	5	Q9VL86	Q9vl86 drosophila
745	5	17.2	396	8	Q02688	Q02688 podospora a	818	5	17.2	430	16	Q9A5U3	Q9a5u3 caulobacter
746	5	17.2	396	16	Q98AR6	Q98ar6 rhizobium l	819	5	17.2	430	16	Q92BW3	Q92bw3 listeria in

820	5	17.2	431	10	Q9FRP2	Q9frp2 oryza sativ	893	5	17.2	476	10	Q9SHQ6	Q9shq6 arabidopsis
821	5	17.2	431	16	Q9X1K0	Q9x1k0 thermotoga	894	5	17.2	476	17	Q97W00	Q97w00 sulfobolus
822	5	17.2	433	2	Q52401	Q52401 pseudomonas	895	5	17.2	477	4	Q9Y6R9	Q9y6r9 homo sapien
823	5	17.2	433	13	Q04099	Q04099 gallus gall	896	5	17.2	478	4	Q9H0J7	Q9h0j7 homo sapien
824	5	17.2	433	16	Q9RZJ8	Q9rzj8 deinococcus	897	5	17.2	478	16	Q9HUZ2	Q9huz2 pseudomonas
825	5	17.2	434	16	Q9KAQ9	Q9kaq9 bacillus ha	898	5	17.2	479	5	Q44505	Q44505 caenorhabdi
826	5	17.2	435	2	Q50522	Q50522 streptomyce	899	5	17.2	479	10	Q9LML6	Q9lml6 arabidopsis
827	5	17.2	435	2	Q9S273	Q9s273 streptomyce	900	5	17.2	480	5	Q22134	Q22134 caenorhabdi
828	5	17.2	435	13	P70033	P70033 xenopus lae	901	5	17.2	481	8	Q9G154	Q9g154 cascaronia
829	5	17.2	436	2	Q45293	Q45293 corynebacte	902	5	17.2	482	2	Q69046	Q69046 rhodococcus
830	5	17.2	436	4	Q9NAB6	Q9nab6 homo sapien	903	5	17.2	482	12	Q91DM3	Q91dm3 turnip mosa
831	5	17.2	437	5	Q24481	Q24481 drosophila	904	5	17.2	483	3	Q96W00	Q96w00 schizosacch
832	5	17.2	437	17	Q9V2J8	Q9v2j8 pyrococcus	905	5	17.2	484	10	Q9C6Y7	Q9c6y7 arabidopsis
833	5	17.2	438	4	Q9NPV9	Q9npv9 homo sapien	906	5	17.2	485	8	Q9G161	Q9g161 kotschy oc
834	5	17.2	439	12	Q72457	Q72457 gloxinia to	907	5	17.2	485	16	Q9C208	Q9c208 rhizobium m
835	5	17.2	439	12	Q84402	Q84402 peanut bud	908	5	17.2	486	16	Q9PNA8	Q9pna8 campylobact
836	5	17.2	439	16	Q9PKM5	Q9pkm5 chlamydia m	909	5	17.2	487	16	P96405	P96405 mycobacteri
837	5	17.2	440	5	Q9W297	Q9w297 drosophila	910	5	17.2	487	17	Q97ZC5	Q97zc5 sulfobolus
838	5	17.2	443	2	Q9L4U5	Q9l4u5 streptomyce	911	5	17.2	492	2	Q51584	Q51584 plectonema
839	5	17.2	443	5	Q9N8D6	Q9n8d6 trypanosoma	912	5	17.2	492	5	Q9NIN9	Q9nin9 babesia mlc
840	5	17.2	443	5	P90628	P90628 leishmania	913	5	17.2	492	8	Q9GHC0	Q9ghc0 streptopus
841	5	17.2	444	2	Q9XD80	Q9xd80 streptomyce	914	5	17.2	492	8	Q9GHC0	Q9ghc0 streptopus
842	5	17.2	444	5	Q9V6N5	Q9v6n5 drosophila	915	5	17.2	494	5	Q04275	Q04275 trypanosoma
843	5	17.2	444	8	Q9TMC3	Q9tmc3 scollopus b	916	5	17.2	495	10	Q9FFM9	Q9ffm9 arabidopsis
844	5	17.2	444	8	Q95DK1	Q95dk1 scollopus h	917	5	17.2	495	17	Q96YR7	Q96yr7 sulfobolus
845	5	17.2	444	11	Q55034	Q55034 rattus norv	918	5	17.2	496	2	Q9AKN4	Q9akn4 rickettsia
846	5	17.2	444	16	Q9I107	Q9i107 pseudomonas	919	5	17.2	496	16	Q987P9	Q987p9 rhizobium l
847	5	17.2	446	13	Q9OYM9	Q9oym9 brachydanio	920	5	17.2	498	4	Q9BUSA	Q9bus4 homo sapien
848	5	17.2	448	5	Q95WG3	Q95wg3 acropora mi	921	5	17.2	499	3	Q9US44	Q9us44 schizosacch
849	5	17.2	448	12	Q9E390	Q9e390 hepatitis c	922	5	17.2	499	5	Q9W0N4	Q9won4 drosophila
850	5	17.2	448	12	Q9E383	Q9e383 hepatitis c	923	5	17.2	499	5	Q44657	Q44657 caenorhabdi
851	5	17.2	448	12	Q9E382	Q9e382 hepatitis c	924	5	17.2	499	5	Q44658	Q44658 caenorhabdi
852	5	17.2	449	16	Q83887	Q83887 treponema p	925	5	17.2	499	5	Q61204	Q61204 caenorhabdi
853	5	17.2	450	3	Q9UTB3	Q9utb3 schizosacch	926	5	17.2	500	3	Q59808	Q59808 schizosacch
854	5	17.2	451	2	Q9AEP6	Q9aep6 vibrio vuln	927	5	17.2	500	4	Q96FE0	Q96fe0 homo sapien
855	5	17.2	451	10	Q82690	Q82690 lepicinolis	928	5	17.2	500	5	Q61935	Q61935 caenorhabdi
856	5	17.2	451	16	Q9KSI0	Q9ksi0 vibrio chol	929	5	17.2	500	5	Q22904	Q22904 caenorhabdi
857	5	17.2	452	16	Q9PAD2	Q9pad2 xylella fas	930	5	17.2	501	16	Q9YH85	Q9ywh85 melanoplus
858	5	17.2	453	2	Q9S202	Q9s202 streptomyce	931	5	17.2	501	16	Q9P089	Q9p089 ureaplasma
859	5	17.2	453	5	Q9V342	Q9v342 drosophila	932	5	17.2	502	16	Q9K107	Q9kl07 neisseria m
860	5	17.2	453	16	Q99V51	Q99v51 staphylococ	933	5	17.2	502	16	Q9JSX4	Q9jsx4 neisseria m
861	5	17.2	454	5	Q9VP00	Q9vp00 drosophila	934	5	17.2	503	5	Q9BLS6	Q9bls6 leishmania
862	5	17.2	456	5	Q9N777	Q9n777 leishmania	935	5	17.2	503	10	Q9AVL5	Q9avl5 panax ginse
863	5	17.2	457	6	Q9XS62	Q9xs62 canis famil	936	5	17.2	503	10	Q9AVL4	Q9avl4 panax quinq
864	5	17.2	457	16	P96742	P96742 bacillus su	937	5	17.2	503	10	Q9AVL3	Q9avl3 panax japon
865	5	17.2	457	16	Q55885	Q55885 synecocyst	938	5	17.2	503	10	Q9AVL2	Q9avl2 panax vietn
866	5	17.2	459	2	Q32384	Q32384 synecococ	939	5	17.2	503	10	Q9AVL1	Q9avl1 panax japon
867	5	17.2	460	2	Q53122	Q53122 rhodococcus	940	5	17.2	503	10	Q9AVL0	Q9avl0 panax pseud
868	5	17.2	460	2	P72256	P72256 rhodococcus	941	5	17.2	504	5	Q95XR7	Q95xr7 caenorhabdi
869	5	17.2	460	4	Q9Y5C1	Q9y5c1 homo sapien	942	5	17.2	504	10	Q40371	Q40371 matteuccia
870	5	17.2	461	12	Q91LR6	Q91lr6 turnip mosa	943	5	17.2	504	11	Q99M43	Q99m43 mus musculu
871	5	17.2	462	2	Q52801	Q52801 amycolatops	944	5	17.2	504	12	Q89807	Q89807 human adeno
872	5	17.2	462	2	Q69823	Q69823 streptomyce	945	5	17.2	505	2	Q9AQ45	Q9aq45 selenomonas
873	5	17.2	463	3	Q9P3C3	Q9p3c3 neurospora	946	5	17.2	506	4	Q9H694	Q9h694 homo sapien
874	5	17.2	463	4	Q9NQ66	Q9nq66 homo sapien	947	5	17.2	506	5	Q26767	Q26767 trypanosoma
875	5	17.2	463	10	Q9FTD2	Q9fid2 arabidopsis	948	5	17.2	506	8	Q47143	Q47143 calluna vul
876	5	17.2	463	16	Q9X143	Q9x143 thermotoga	949	5	17.2	506	10	Q49373	Q49373 arabidopsis
877	5	17.2	465	2	Q68123	Q68123 rhodobacter	950	5	17.2	506	10	Q9CA74	Q9ca74 arabidopsis
878	5	17.2	465	5	Q9VAK4	Q9vak4 drosophila	951	5	17.2	507	2	Q9K3F2	Q9k3f2 streptomyce
879	5	17.2	465	10	Q9FW59	Q9fw59 oryza sativ	952	5	17.2	507	15	Q88142	Q88142 chimpanzee
880	5	17.2	467	3	Q93838	Q93838 aspergillus	953	5	17.2	507	17	Q26749	Q26749 methanother
881	5	17.2	467	3	Q9U0Z7	Q9u0z7 aspergillus	954	5	17.2	508	12	Q9QAH8	Q9qah8 human adeno
882	5	17.2	467	3	Q9HEQ0	Q9heq0 aspergillus	955	5	17.2	508	13	Q9YHV9	Q9yvh9 gallus gall
883	5	17.2	467	12	Q9EA46	Q9ea46 human adeno	956	5	17.2	508	13	Q9PVK0	Q9pvk0 gallus gall
884	5	17.2	468	16	Q9CMC2	Q9cmc2 pasteurella	957	5	17.2	509	16	Q9CMQ9	Q9cmq9 pasteurella
885	5	17.2	469	10	Q9SVA8	Q9sva8 arabidopsis	958	5	17.2	509	17	Q9HNZ0	Q9hnz0 halobacteri
886	5	17.2	470	16	Q9AB40	Q9ab40 caulobacter	959	5	17.2	509	17	Q9HMH4	Q9hnh4 halobacteri
887	5	17.2	472	17	Q972D2	Q972d2 sulfobolus	960	5	17.2	510	8	Q9TK13	Q9tk13 aeshchynomen
888	5	17.2	473	10	Q22562	Q22562 arabidopsis	961	5	17.2	510	16	Q92SN2	Q92sn2 rhizobium m
889	5	17.2	474	2	Q33654	Q33654 streptococ	962	5	17.2	511	8	Q9TK12	Q9tk12 pterocarpus
890	5	17.2	474	5	Q9XXQ2	Q9xxq2 caenorhabdi	963	5	17.2	511	8	Q9GHF9	Q9ghf9 aletris gla
891	5	17.2	474	11	Q9ESD0	Q9esd0 mus musculu	964	5	17.2	511	8	Q9GHF8	Q9ghf8 aletris gla
892	5	17.2	476	6	Q9GMV3	Q9gmv3 macaca fasc	965	5	17.2	511	8	Q9GHF0	Q9ghf0 aletris lut

966	5	17.2	511	8	Q9GI70	Q9gi170 inocarpus f	1039	5	17.2	525	10	O80664	O80664 arabidopsis
967	5	17.2	511	16	Q9A4C3	Q9a4c3 caulobacter	1040	5	17.2	526	2	O68660	O68660 actinobacill
968	5	17.2	512	8	Q9TKT3	Q9tkt3 pterocarpus	1041	5	17.2	526	2	P95663	P95663 rickettsia
969	5	17.2	512	8	Q9TK14	Q9tk14 aescynomen	1042	5	17.2	526	5	O44659	O44659 caenorhabdi
970	5	17.2	512	8	Q9GI86	Q9gi86 grazieloden	1043	5	17.2	526	5	Q9N9B7	Q9n9b7 leishmania
971	5	17.2	512	8	Q9GI79	Q9gi79 aescynomen	1044	5	17.2	526	5	Q95PT4	Q95pt4 leishmania
972	5	17.2	512	8	Q9GI77	Q9gi77 platymisciu	1045	5	17.2	526	17	Q9UZ04	Q9uz04 pyrococcus
973	5	17.2	512	8	Q9GI76	Q9gi76 platymisciu	1046	5	17.2	527	5	O00815	O00815 leishmania
974	5	17.2	512	8	Q9GI75	Q9gi75 discolobium	1047	5	17.2	527	10	Q94A60	Q94a60 arabidopsis
975	5	17.2	512	8	Q9GI74	Q9gi74 discolobium	1048	5	17.2	527	2	O31305	O31305 borrelia bu
976	5	17.2	512	8	Q9GI69	Q9gi69 geofiroea s	1049	5	17.2	528	3	Q9P7P2	Q9p7p2 schizosacch
977	5	17.2	512	8	Q9GI68	Q9gi68 geofiroea d	1050	5	17.2	528	4	Q9H7D6	Q9h7d6 homo sapien
978	5	17.2	512	8	Q9GI67	Q9gi67 ramorinoa g	1051	5	17.2	528	10	Q9C880	Q9c880 arabidopsis
979	5	17.2	512	8	Q9GI66	Q9gi66 tipuana tip	1052	5	17.2	528	16	O54584	O54584 borrelia bu
980	5	17.2	512	8	Q9GI65	Q9gi65 centrolobiu	1053	5	17.2	528	16	Q9RTP5	Q9rtp5 deinococcus
981	5	17.2	512	8	Q9GD26	Q9gd26 aescynomen	1054	5	17.2	529	17	Q976T0	Q976t0 sulfolobus
982	5	17.2	512	9	Q9XJG2	Q9xjg2 bacterioph	1055	5	17.2	530	13	Q9W629	Q9w629 brachydanio
983	5	17.2	514	5	O62196	O62196 caenorhabdi	1056	5	17.2	531	5	Q9V5M3	Q9v5m3 drosophila
984	5	17.2	514	8	Q9TK31	Q9tk31 ormocarpops	1057	5	17.2	531	10	Q942D9	Q942d9 oryza sativ
985	5	17.2	514	8	Q9TK30	Q9tk30 ormocarpops	1058	5	17.2	532	10	O81849	O81849 arabidopsis
986	5	17.2	514	8	Q9TK29	Q9tk29 ormocarpum	1059	5	17.2	532	12	O65871	O65871 barley yell
987	5	17.2	514	8	Q9TK28	Q9tk28 ormocarpum	1060	5	17.2	532	12	O65876	O65876 barley yell
988	5	17.2	514	8	Q9TK27	Q9tk27 ormocarpum	1061	5	17.2	532	16	Q92R32	Q92r32 rhizobium m
989	5	17.2	514	8	Q9TK26	Q9tk26 ormocarpum	1062	5	17.2	533	4	O94968	O94968 homo sapien
990	5	17.2	514	8	Q9TK25	Q9tk25 ormocarpum	1063	5	17.2	534	10	Q9LG16	Q9lg16 oryza sativ
991	5	17.2	514	8	Q9TK24	Q9tk24 diphsa ame	1064	5	17.2	535	16	Q9ABT0	Q9abt0 caulobacter
992	5	17.2	514	8	Q9TK23	Q9tk23 diphsa flo	1065	5	17.2	537	16	Q9CD55	Q9cd55 mycobacteri
993	5	17.2	514	8	Q9TK22	Q9tk22 diphsa thu	1066	5	17.2	538	2	O68977	O68977 sphingomona
994	5	17.2	514	8	Q9TK21	Q9tk21 pictetia ac	1067	5	17.2	538	2	O53961	O53961 flavobacter
995	5	17.2	514	8	Q9TK20	Q9tk20 pictetia ma	1068	5	17.2	538	5	O77100	O77100 aedes aegyp
996	5	17.2	514	8	Q9TK19	Q9tk19 pictetia an	1069	5	17.2	538	16	Q55562	Q55562 synecocyst
997	5	17.2	514	8	Q9TK01	Q9tk01 diphsa orm	1070	5	17.2	539	4	O60774	O60774 homo sapien
998	5	17.2	514	8	Q9GI72	Q9gi72 cranocarpus	1071	5	17.2	539	10	Q943P8	Q943p8 oryza sativ
999	5	17.2	514	8	Q9GI73	Q9gi73 brva ebenus	1072	5	17.2	541	5	Q9VJL5	Q9vjl5 drosophila
1000	5	17.2	514	8	Q9GI71	Q9gi71 platypodium	1073	5	17.2	542	6	Q9N280	Q9n280 canis famil
1001	5	17.2	514	12	Q9JEV9	Q9jev9 human adeno	1074	5	17.2	542	17	Q980L0	Q980l0 sulfolobus
1002	5	17.2	514	12	Q9EA43	Q9ea43 human adeno	1075	5	17.2	544	2	Q9FCG8	Q9fcg8 streptomyce
1003	5	17.2	514	16	Q982F3	Q982f3 rhizobium l	1076	5	17.2	544	4	O13652	O13652 homo sapien
1004	5	17.2	515	8	Q9GHD7	Q9ghd7 heloniopsis	1077	5	17.2	545	3	O59836	O59836 schizosacch
1005	5	17.2	515	8	Q9GHD6	Q9ghd6 heloniopsis	1078	5	17.2	545	3	Q9HEK1	Q9hek1 neurospora
1006	5	17.2	515	8	Q9GHD5	Q9ghd5 heloniopsis	1079	5	17.2	545	5	Q9TXQ5	Q9txq5 caenorhabdi
1007	5	17.2	515	8	Q9GHD4	Q9ghd4 heloniopsis	1080	5	17.2	546	4	Q9NRG9	Q9nrg9 homo sapien
1008	5	17.2	515	8	Q9GHD3	Q9ghd3 heloniopsis	1081	5	17.2	546	4	Q9H5T4	Q9h5t4 homo sapien
1009	5	17.2	515	8	Q9GHD2	Q9ghd2 heloniopsis	1082	5	17.2	546	4	Q9NMW6	Q9nmw6 homo sapien
1010	5	17.2	515	8	Q9GHD1	Q9ghd1 heloniopsis	1083	5	17.2	550	5	Q9XVE3	Q9xve3 caenorhabdi
1011	5	17.2	515	8	Q9GHD0	Q9ghd0 heloniopsis	1084	5	17.2	550	10	O91LK00	O91lk00 arabidopsis
1012	5	17.2	515	8	Q9GHC9	Q9ghc9 heloniopsis	1085	5	17.2	550	16	Q9KPA4	Q9kpa4 vibrio chol
1013	5	17.2	515	8	Q9GHC8	Q9ghc8 heloniopsis	1086	5	17.2	551	16	Q99VA0	Q99va0 staphylococ
1014	5	17.2	515	12	Q9JEV8	Q9jev8 human adeno	1087	5	17.2	551	16	Q93VD5	Q93vd5 oryza sativ
1015	5	17.2	515	12	Q9EA44	Q9ea44 human adeno	1088	5	17.2	553	5	Q9XD30	Q9xd30 caenorhabdi
1016	5	17.2	516	3	Q9P775	Q9p775 schizosacch	1089	5	17.2	554	5	Q95T78	Q95t78 drosophila
1017	5	17.2	517	4	Q9V202	Q9v2q2 homo sapien	1090	5	17.2	554	10	O9AU55	O9au55 asparagus f
1018	5	17.2	517	4	Q9N051	Q9nc51 homo sapien	1091	5	17.2	554	10	Q93VF8	Q93vf8 oryza sativ
1019	5	17.2	517	12	Q9WF17	Q9wfl17 human adeno	1092	5	17.2	561	4	Q9HA37	Q9ha37 homo sapien
1020	5	17.2	517	12	Q91J05	Q91j05 human adeno	1093	5	17.2	561	4	Q96SK2	Q96sk2 homo sapien
1021	5	17.2	517	12	Q91J00	Q91j00 human adeno	1094	5	17.2	564	5	Q9VTS7	Q9vts7 drosophila
1022	5	17.2	518	2	Q9EZ54	Q9ez54 arthrobacte	1095	5	17.2	564	5	Q95U49	Q95u49 drosophila
1023	5	17.2	519	12	Q9JEW0	Q9jew0 human adeno	1096	5	17.2	564	16	O9PAD6	O9pad6 xyella fas
1024	5	17.2	519	12	Q9EA45	Q9ea45 human adeno	1097	5	17.2	565	5	O77099	O77099 aedes aegyp
1025	5	17.2	519	12	Q91J02	Q91j02 human adeno	1098	5	17.2	567	5	Q9NK58	Q9nk58 drosophila
1026	5	17.2	519	12	Q91J01	Q91j01 human adeno	1099	5	17.2	567	10	Q9ZQN5	Q9zqn5 arabidopsis
1027	5	17.2	520	2	Q9AJQ5	Q9ajq5 arthrobacte	1100	5	17.2	568	5	O16192	O16192 caenorhabdi
1028	5	17.2	520	2	Q9AJQ3	Q9ajq3 arthrobacte	1101	5	17.2	569	13	O57339	O57339 xenopus lae
1029	5	17.2	520	12	Q91W62	Q91w62 human adeno	1102	5	17.2	569	16	O9CCV1	O9ccv1 mycobacteri
1030	5	17.2	520	12	Q91J03	Q91j03 human adeno	1103	5	17.2	570	5	O44736	O44736 caenorhabdi
1031	5	17.2	520	17	O58095	O58095 pyrococcus	1104	5	17.2	570	5	O95SD1	O95sd1 drosophila
1032	5	17.2	522	2	O85077	O85077 arthrobacte	1105	5	17.2	571	12	Q910Z6	Q910z6 human adeno
1033	5	17.2	522	2	O44414	O44414 arthrobacte	1106	5	17.2	571	12	Q910Z7	Q910z7 human adeno
1034	5	17.2	522	12	Q912J4	Q912j4 human adeno	1107	5	17.2	571	12	O910U2	O910u2 human adeno
1035	5	17.2	523	10	O9C8V6	O9c8v6 arabidopsis	1108	5	17.2	571	16	O66855	O66855 aquifex aco
1036	5	17.2	524	10	O9FUJ1	O9fuj1 arabidopsis	1109	5	17.2	575	12	O11417	O11417 duck adenov
1037	5	17.2	525	2	O50903	O50903 myxococcus	1110	5	17.2	575	16	P73479	P73479 synecocyst
1038	5	17.2	525	4	Q96KC6	Q96kc6 homo sapien	1111	5	17.2	575	16	Q9KA17	Q9ka17 bacillus ha

DR EMBL; AB032826; BAA84781.1; -.
SQ SEQUENCE 386 AA; 43061 MW; EC6F3A76874E528F CRC64;

Query Match 31.0%; Score 9; DB 6; Length 386;
Best Local Similarity 100.0%; Pred. No. 0.053; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRCINQLLC 9
|||||||

Db 283 KRCINQLLC 291

RESULT 3

Q9E1T3 ID Q9E1T3 PRELIMINARY; PRT; 28 AA.

AC Q9E1T3

DT 01-MAR-2001 (TReMBLrel. 16, Created)

DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)

DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)

DE L PROTEIN (FRAGMENT).

OS rabies virus.

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;

OC Rhabdoviridae; Lyssavirus.

OX NCBI_TaxID=11292;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=39481;

RA Paez A., Duarte X., Pulido S.;

RT "Molecular epidemiology of rabies epizootics in Colombia.";

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

RE EMBL; AF276906; AAG24801.1; -. 28

FT NON_TER 28

SQ SEQUENCE 28 AA; 2953 MW; 2E330EA9AB8F3C6A CRC64;

Query Match 24.1%; Score 7; DB 12; Length 28;

Best Local Similarity 100.0%; Pred. No. 0.78;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 DSELAPR 20

|||||||

Db 16 DSELAPR 22

RESULT 4

Q06858 ID Q06858 PRELIMINARY; PRT; 495 AA.

AC Q06858

DT 01-NOV-1996 (TReMBLrel. 01, Created)

DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)

DE PHO2 PROTEIN.

GN PHO2 OR YDL106C.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=GG100-14D;

RA Bergman L.W.;

RL Submitted (AUG-1990) to the EMBL/GenBank/DBJ databases.

RE [2]

RP SEQUENCE OF 1-184 FROM N.A.

RC STRAIN=GG100-14D;

RX MEDLINE=91288241; PubMed=2062666;

RA McCarthy B.J., Creasy C.L., Bergman L.W.;

RT "Molecular analysis of a temperature sensitive allele of the PHO2 gene

of Saccharomyces cerevisiae.";

RL Nucleic Acids Res. 19; 3463-3463(1991).

CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

CC -1- SIMILARITY: WITH OTHER HOMEBOX PROTEINS.

DR EMBL; X54293; CAA38192.1; -.
DR SGD; S0002264; PHO2.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
SQ SEQUENCE 495 AA; 55527 MW; 781D3A5B0E6E6F2D CRC64;

Query Match 24.1%; Score 7; DB 3; Length 495;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 LPTDSEL 17

|||||||

Db 469 LPTDSEL 475

RESULT 5

Q9M295 ID Q9M295 PRELIMINARY; PRT; 1009 AA.

AC Q9M295

DT 01-OCT-2000 (TReMBLrel. 15, Created)

DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)

DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)

DE HYPOTHETICAL 113.2 KDA PROTEIN.

GN T12K4.140.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Monfort A., Casacuberta E., Puldomenech P., Mewes H.W., Rudd S.,

RL Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

RE [2]

RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; ALI38640; CAB86452.1; -. 1

DR InterPro; IPR003653; SUMO_protease.

KW Hypothetical protein.

SQ SEQUENCE 1009 AA; 113151 MW; 816FAF14110588A7 CRC64;

Query Match 24.1%; Score 7; DB 10; Length 1009;

Best Local Similarity 100.0%; Pred. No. 18;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KLPTDSE 16

|||||||

Db 82 KLPTDSE 88

RESULT 6

Q42306 ID Q42306 PRELIMINARY; PRT; 98 AA.

AC Q42306

DT 01-NOV-1996 (TReMBLrel. 01, Created)

DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)

DE RIBULOSE BISPHOSPHATE CARBOXYLASE (FRAGMENT).

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-CELL SUSPENSION CULTURE OF ECOTYPE COLUMBIA;
 RA Barlett C., Dabos P., Tremousaygue D., Lescure B.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z37217; CA85525.1; -.
 DR InterPro; IPR000894; Rubisco_small.
 DR Pfam; PF00101; Rubisco_small; 1.
 DR PRINTS; PR00152; RUBISCO SMALL.
 DR ProDom; PD000290; RUBISCO_small; 1.
 FT NON_TER 98
 FT SEQUENCE 98 AA; 10793 MW; 79A95053334ACFD1 CRC64;

Query Match 20.7%; Score 6; DB 10; Length 98;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDSELA 18
 Db 73 TDSELA 78

RESULT 7
 Q9JUF2 PRELIMINARY; PRT; 103 AA.
 AC Q9JUF2;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE BRAIN CDNA, CLONE MNCB-0953, SIMILAR TO MUS MUSCULUS SPAG4 PROTEIN
 DE ALTERNATIVE SPLICED.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL;
 RA Osada N., Kusuda J., Tanuma R., Ito A., Hirata M., Sugano S.,
 RA Hashimoto K.;
 RT "Isolation of full-length cDNA clones from mouse brain cDNA library
 RT made by oligo-capping method.";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB041554; BAA95039.1; -.
 SQ SEQUENCE 103 AA; 10590 MW; E8707EEA59ED8372 CRC64;

Query Match 20.7%; Score 6; DB 11; Length 103;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELAPRS 21
 Db 81 ELAPRS 86

RESULT 8
 Q42046 PRELIMINARY; PRT; 107 AA.
 AC Q42046;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE RIBULOSE BISPHOSPHATE CARBOXYLASE (FRAGMENT).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA; TISSUE=SEEDLING;
 RA Hofte H.;
 RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.

DR EMBL; Z25565; CA80994.1; -.
 DR InterPro; IPR000894; Rubisco_small.
 DR Pfam; PF00101; Rubisco_small; 1.
 DR PRINTS; PR00152; RUBISCO SMALL.
 DR ProDom; PD000290; RUBISCO_small; 1.
 FT NON_TER 107
 FT SEQUENCE 107 AA; 11891 MW; 30018913B3C8971B CRC64;

Query Match 20.7%; Score 6; DB 10; Length 107;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDSELA 18
 Db 76 TDSELA 81

RESULT 9
 Q94JW2 PRELIMINARY; PRT; 125 AA.
 AC Q94JW2;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE FLO19.10/FLO19.10.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Shinn P., Banh J.,
 RA Bowsher L., Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
 RA Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M.,
 RA Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M.,
 RA Seki M., Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K.,
 RA Davis R.W., Theologis A., Ecker J.R.;
 RT "Arabidopsis cDNA clones.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF372874; AAK49590.1; -.
 SQ SEQUENCE 125 AA; 14699 MW; 7173FEEF0E1A2FCC CRC64;

Query Match 20.7%; Score 6; DB 10; Length 125;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDSELA 18
 Db 22 TDSELA 27

RESULT 10
 Q32096 PRELIMINARY; PRT; 129 AA.
 AC Q32096;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
 DE YUZE PROTEIN.
 GN YUZE.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Broutillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.-K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaere-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takenaru K.,
 RA Taseuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
 RA "The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*.";
 RT Nature 390:249-256(1997).
 RL Nature 390:249-256(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; 299120; CAB15169.1; -;
 KW Complete proteome.
 SQ SEQUENCE 129 AA; 14752 MW; 25F6B7F09069DCE7 CRC64;

Query Match 20.7%; Score 6; DB 16; Length 129;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDSELA 18
 |||||
 Db 8 TDSELA 13

RESULT 11
 O36891 PRELIMINARY; PRT; 135 AA.
 ID O36891
 AC O36891;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DR 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE GAG POLYPROTEIN (FRAGMENT).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97445059; PubMed=9300048;
 RA Leigh Brown A.J., Lobdel D., Wade C.M., Rebus S., Phillips N.,
 RA Brettie R.P., France A.J., Leen C.S., McMenamin J., McMillan A.,
 RA Maw R.D., Mulcahy F., Robertson J.R., Sankar K.N., Scott G., Wyld R.,
 RA Petherer J.F.;
 RT "The molecular epidemiology of human immunodeficiency virus type 1 in
 six cities in Britain and Ireland.";
 RL Virology 235:166-177(1997).
 DR EMBL; AF014288; AAC58369.1; -;

DR InterPro: IPR000071; Retroviral_gag_p17.
 DR Pfam: PF00540; gag_p17; 1.
 DR PRINTS; PR00234; HIVMATRIX.
 KW AIDS; Core protein; Polyprotein.
 FT NON_TER 1
 FT NON_TER 135
 SQ SEQUENCE 135 AA; 15010 MW; BC9DD8D708D01F44 CRC64;
 Query Match 20.7%; Score 6; DB 15; Length 135;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 11 LPTDSE 16
 |||||
 Db 45 LPTDSE 50
 RESULT 12
 Q9Y874 PRELIMINARY; PRT; 167 AA.
 ID Q9Y874
 AC Q9Y874;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DR 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE V-TYPE ATPASE SUBUNIT C' (VACUOLAR ATPASE PROTEOLIPID SUBUNIT
 DE C').
 GN VMA-11.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-OAKRIDGE 74;
 RA Pounder J.I., Bowman B.J.;
 RT "Proteolipid subunits of the vacuolar ATPase in *Neurospora crassa*.";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Tenney K.L., Bowman B.J.;
 RT "Genomic sequence of the c' subunit of the vacuolar ATPase.";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF162776; AAD45120.2; -;
 DR EMBL; AF330696; AAK13465.1; -;
 DR InterPro: IPR002379; ATPase_C.
 DR InterPro: IPR000245; Vac_ATPase.
 DR Pfam: PF00137; ATP-synt_C; 2.
 DR PRINTS; PR00122; VACATPASE.
 SQ SEQUENCE 167 AA; 17029 MW; F30C41CF2E277745 CRC64;

Query Match 20.7%; Score 6; DB 3; Length 167;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 DSELAP 19
 |||||
 Db 7 DSELAP 12

RESULT 13
 Q9FP16 PRELIMINARY; PRT; 180 AA.
 ID Q9FP16
 AC Q9FP16;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DR 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE ATIG67090.
 GN ATIG67090/F1019.10.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;

OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Shinn P., Chao Q., Brooks S., Chen H., Cheuk R., Johnson-Hopson C.,
 RA Khan S., Kim C.J., Banh J., Bowser L., Chung M.K., Goldsmith A.D.,
 RA Jones T., Karlin-Neumann G., Lam B., Lee J.M., Lin J., Liu S.X.,
 RA Miranda M., Nguyen M., Palm C.J., Pham P.K., Quach H.L., Sakano H.,
 RA Southwick A., Toriumi M., Yamada K., Yu G., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RA "Arabidopsis cDNA clones";
 RT Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AF325004; AAG40356.1; -;
 DR InterPro; IPR000894; RUBISCO_small.
 DR Pfam; PF00101; RUBISCO_small; 1.
 DR PRINTS; PR00152; RUBISCO_SMALL.
 DR ProDom; PD000290; RUBISCO_small; 1.
 SQ SEQUENCE 180 AA; 20315 MW; 5DFB0417710D98F7 CRC64;

Query Match 20.7%; Score 6; DB 10; Length 180;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 TDSELA 18
 Db |||||
 77 TDSELA 82

RESULT 14

ID Q9SAV4 PRELIMINARY; PRT; 180 AA.
 AC Q9SAV4;
 DT 01-MAY-2000 (TREMELrel. 13, Created)
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)
 DE RIBULOSE BISPHOSPHATE CARBOXYLASE, SMALL SUBUNIT (EC 4.1.1.39)
 DE (00010C11) (F1019.10/F1019.10) (AT1G67090/F1019.10).
 GN F5A8.1 OR AT1G67090.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Federspiel N.A., Palm C.J., Conway A.B., Kurtz D.B., Conway A.R.,
 RA Au M., Araujo R., Buehler E., Dewar K., Feng J., Kim C., Li Y.,
 RA Oji O., Osborne B.I., Shinn P., Sun H., Toriumi M., Vysotskaia V.S.,
 RA Yu G., Ecker J., Theologis A., Davis R.W.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]

RP SEQUENCE FROM N.A.
 RA Chao Q., Brooks S., Chen H., Johnson-Hopson C., Khan S., Shinn P.,
 RA Ecker J.;
 RT "Full length cDNA sequence of Arabidopsis thaliana";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP SEQUENCE FROM N.A.
 RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]

RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,

RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 RN [6]

RP SEQUENCE FROM N.A.
 RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
 RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
 RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
 RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
 RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 RT "Arabidopsis cDNA clones";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC004146; AAD10655.1; -;
 DR EMBL; AF325011; AAG40363.1; -;
 DR EMBL; AY059940; AAL24422.1; -;
 DR EMBL; AF410291; AAK95277.1; -;
 DR EMBL; AY054581; AAK96772.1; -;
 DR EMBL; AY054188; AAL06849.1; -;
 DR EMBL; AY059831; AAL24219.1; -;
 DR InterPro; IPR000894; RUBISCO_small.
 DR Pfam; PF00101; RUBISCO_small; 1.
 DR PRINTS; PR00152; RUBISCO_SMALL.
 DR ProDom; PD000290; RUBISCO_small; 1.
 KW Lyase.

SQ SEQUENCE 180 AA; 20216 MW; 5DE768676D61F8F7 CRC64;
 Query Match 20.7%; Score 6; DB 10; Length 180;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 TDSELA 18
 Db |||||
 77 TDSELA 82

RESULT 15

ID Q9X903 PRELIMINARY; PRT; 207 AA.
 AC Q9X903;
 DT 01-NOV-1999 (TREMELrel. 12, Created)
 DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)
 DE PUTATIVE CARBONIC ANHYDRASE.
 GN SCH35.03.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Oliver K., Harris D.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kieser H.M., Denaparte D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL; AL078610; CAB44424.1; -;
 DR InterPro; IPR001765; Prok_CO2_anhydrase.
 DR Pfam; PF00484; Pro_CA; 1.
 DR PROSITE; PS00704; PROK_CO2_ANHYDRASE.1; UNKNOWN.1.
 SQ SEQUENCE 207 AA; 21442 MW; 4A9CBA8926018568 CRC64;

Query Match 20.7%; Score 6; DB 2; Length 207;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SELAPR 20
 |||||
 DB 39 SELAPR 44

RESULT 16
 Q9L0K0
 ID Q9L0K0 PRELIMINARY; PRT; 209 AA.
 AC Q9L0K0;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
 DE PUTATIVE SECRETED PROTEIN.
 GN SCD40A.11C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Seeger K.J., Harris D.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kieser H.M., Denaparte D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL; AL161691; CAB81856.1; -;
 SQ SEQUENCE 209 AA; 21951 MW; 63BB4EEFB9718BC CRC64;

Query Match 20.7%; Score 6; DB 2; Length 209;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELAPRS 21
 |||||
 DB 187 ELAPRS 192

RESULT 17
 Q9FGE8

ID Q9FGE8 PRELIMINARY; PRT; 212 AA.
 AC Q9FGE8;
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
 DE GENOMIC DNA, CHROMOSOME 5, BAC CLONE:T12B11.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI.";
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB025640; BAB11609.1; -;
 SQ SEQUENCE 212 AA; 24693 MW; F45BCB218FFDB488 CRC64;

Query Match 20.7%; Score 6; DB 10; Length 212;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDSELA 18
 |||||
 DB 89 TDSELA 94

RESULT 18
 Q18836
 ID Q18836 PRELIMINARY; PRT; 228 AA.
 AC Q18836; Q19856;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE C54D10.10 PROTEIN.
 GN C54D10.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kershaw J.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94150718; PubMed=7906398;
 RA Willson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
 RA Saldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans.";
 RN Nature 368:32-38(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Dobson R.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z75539; CAA99845.1; -;
 DR EMBL; Z75531; CAA99845.1; JOINED.
 DR EMBL; Z75531; CAA99808.1; -;
 DR EMBL; Z75539; CAA99808.1; JOINED.
 DR HSSP; P12111; 2KNT.

DR InterPro; IPR002899; EB.
DR InterPro; IPR002223; Kunitz_BPTI.
DR Pfam; PF00014; Kunitz_BPTI; 2.
DR PRINTS; PR00759; BASICPTASE.
DR SMART; SM00131; KU; 2.
DR SMART; SM00289; WRI; 1.
DR PROSITE; PS00279; BPTI_KUNITZ_2; 2.
KW Serine protease inhibitor.
SQ SEQUENCE 228 AA; 25884 MW; B7C8C1FB71CC06DB CRC64;

Query Match 20.7%; Score 6; DB 5; Length 228;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 CKLPTD 14
| | | | |
DB 133 CKLPTD 138

RESULT 19
O99882

ID O99882 PRELIMINARY; PRT; 246 AA.
AC O99882;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE RNA POLYMERASE (FRAGMENT).
OS Agaricus subrutilescens (Wine-colored agaricus).
OG Mitochondrion. Plasmid pEM.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Agaricaceae; Agaricus.
OX NCBI_TaxID=87253;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-RWK1493;
RX MEDLINE=99172296; PubMed=10072320;
RA Robison M.M., Horgen P.A.;
RT "Widespread distribution of low-copy-number variants of mitochondrial plasmid pEM in the genus Agaricus."
RL Fungal Genet. Biol. 26:62-70(1999).
DR EMBL; AF096909; AAD11786.1; -.
DR HSSP; P00573; IARO.
DR InterPro; IPR002092; RNA_pol_phase.
DR Pfam; PF00940; RNA_pol_1.
DR PROSITE; PS00900; RNA_POL_PHASE_1; 1.
DR PROSITE; PS00489; RNA_POL_PHASE_2; 1.
KW Plasmid.
FT NON_TER 1
FT NON_TER 246
SQ SEQUENCE 246 AA; 28049 MW; 031748DDA3B2AF78 CRC64;

Query Match 20.7%; Score 6; DB 8; Length 246;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KLPTDS 15
| | | | |
DB 237 KLPTDS 242

RESULT 20
O99883

ID O99883 PRELIMINARY; PRT; 246 AA.
AC O99883;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE RNA POLYMERASE (FRAGMENT).
OS Agaricus crocodilinus.
OG Mitochondrion. Plasmid pEM.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;

OC Agaricales; Agaricaceae; Agaricus.
OX NCBI_TaxID=87250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RWK1467;
RX MEDLINE=99172296; PubMed=10072320;
RA Robison M.M., Horgen P.A.;
RT "Widespread distribution of low-copy-number variants of mitochondrial plasmid pEM in the genus Agaricus."
RL Fungal Genet. Biol. 26:62-70(1999).
DR EMBL; AF096912; AAD11787.1; -.
DR HSSP; P00573; IARO.
DR InterPro; IPR002092; RNA_pol_phase.
DR Pfam; PF00940; RNA_pol_1.
DR PROSITE; PS00900; RNA_POL_PHASE_1; 1.
DR PROSITE; PS00489; RNA_POL_PHASE_2; 1.
KW Plasmid.
FT NON_TER 1
FT NON_TER 246
SQ SEQUENCE 246 AA; 28011 MW; AC4883BFFEBB31BF CRC64;

Query Match 20.7%; Score 6; DB 8; Length 246;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KLPTDS 15
| | | | |
DB 237 KLPTDS 242

RESULT 21
O9TFX8

ID O9TFX8 PRELIMINARY; PRT; 246 AA.
AC O9TFX8;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE RNA POLYMERASE (FRAGMENT).
OS Agaricus bitorquis.
OG Mitochondrion. Plasmid pEM.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Agaricaceae; Agaricus.
OX NCBI_TaxID=5343;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-AG17;
RX MEDLINE=99172296; PubMed=10072320;
RA Robison M.M., Horgen P.A.;
RT "Widespread distribution of low-copy-number variants of mitochondrial plasmid pEM in the genus Agaricus."
RL Fungal Genet. Biol. 26:62-70(1999).
DR EMBL; AF096908; AAD11785.1; -.
DR HSSP; P00573; IARO.
DR InterPro; IPR002092; RNA_pol_phase.
DR Pfam; PF00940; RNA_pol_1.
DR PROSITE; PS00900; RNA_POL_PHASE_1; 1.
DR PROSITE; PS00489; RNA_POL_PHASE_2; 1.
KW Plasmid.
FT NON_TER 1
FT NON_TER 246
SQ SEQUENCE 246 AA; 28093 MW; AD5C3CE29633AEC2 CRC64;

Query Match 20.7%; Score 6; DB 8; Length 246;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KLPTDS 15
| | | | |
DB 237 KLPTDS 242

```

RESULT 22
Q9XDW2 ID Q9XDW2 PRELIMINARY; PRT; 248 AA.
AC Q9XDW2;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PSBI.
GN PSBI.
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group; Rhodopseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NO.7.;
RX MEDLINE=20358711; PubMed=10902905;
RA Puskas L.G., Inui M., Kele Z., Yukawa H.;
RT "Cloning of genes participating in aerobic biodegradation of p-cumate
from Rhodopseudomonas palustris.";
RL DNA Seq. 11:9-20(2000).
CC -1- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL: AB022919; BAA82120.1; -.
DR InterPro: IPR001647; HTH_Tetr.
DR Pfam: PF00440; tetr; 1.
DR PRINTS; PR00455; HTHTETR.
KW DNA-binding; Transcription regulation.
SQ SEQUENCE 248 AA; 27742 MW; 39BE352EC3C4F755 CRC64;

Query Match 20.7%; Score 6; DB 2; Length 248;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 DSELAP 19
Db 236 DSELAP 241

RESULT 23
Q931A2 ID Q931A2 PRELIMINARY; PRT; 259 AA.
AC Q931A2;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PUTATIVE ABC TRANSPORTER ATP-BINDING PROTEIN.
GN SMA0083.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymA (megaplasmid 1).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396509; PubMed=11481432;
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA Barlett-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
Sinorhizobium meliloti pSymA megaplasmid.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
DR EMBL; AE007198; AAK64700.1; -.
KW ATP-binding; Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 259 AA; 28810 MW; 4CD2F11AC63F8435 CRC64;

Query Match 20.7%; Score 6; DB 16; Length 259;
Best Local Similarity 100.0%; Pred. No. 67;

```

```

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RCINQL 7
Db 64 RCINQL 69

RESULT 24
Q9DAU1 ID Q9DAU1 PRELIMINARY; PRT; 276 AA.
AC Q9DAU1;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE 1600025D17RIK PROTEIN (PUTATIVE RETINOIC ACID-REGULATED PROTEIN)
GN (RIKEN CDNA 1600025D17 GENE).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PLACENTA;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirai L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Glozak M.A., Li Y., Reuille R., Rogers M.B.;
RT "Trapping and characterization of novel retinoic acid response
elements.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY, NORMAL. 5 MONTH OLD MALE MOUSE.;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK055332; BAB24103.1; -.
DR EMBL; AF361644; AAK52494.1; -.
DR EMBL; BC013549; AAH13549.1; -.
DR MGD; MGI:1919279; 1600025D17Rik.
SQ SEQUENCE 276 AA; 30538 MW; 7C69F60BA3BE1745 CRC64;

Query Match 20.7%; Score 6; DB 11; Length 276;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SELAPR 20
Db 5 SELAPR 10

RESULT 25

```

```
Q9KPY3
ID Q9KPY3 PRELIMINARY; PRT; 281 AA.
AC Q9KPY3
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN VC2229.
GN VC2229.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
[1]
SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberger J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
DR EMBL; AE004294; AAF95373.1; -.
DR TIGR; VC2229; -.
KW Complete proteome.
SQ SEQUENCE 281 AA; 31824 MW; 7CA75AD3494DFDD0 CRC64;

Query Match 20.7%; Score 6; DB 16; Length 281;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDSELA 18
Db 128 TDSELA 133

RESULT 26
Q92PN5
ID Q92PN5 PRELIMINARY; PRT; 295 AA.
AC Q92PN5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PUTATIVE OXIDOREDUCTASE PROTEIN.
GN SMC00274.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
[1]
SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21368234; PubMed=11474104;
RA Galibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F.,
RA Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G.,
RA Boutry M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P.,
RA Cowie A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F.,
RA Gloux S., Godrie T., Goffeau A., Golding B., Gouz J., Gurjal M.,
RA Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,
RA Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Lelaure V.,
RA Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B.,
RA Ramsperger U., Surzycki R., Thebault P., Vandenbol M.,
RA Vorhoefer F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;
RT "The composite genome of the legume symbiont Sinorhizobium meliloti."
RL Science 293:668-672(2001).
DR EMBL; AL591788; CAC46281.1; -.
KW Complete proteome.
SQ SEQUENCE 295 AA; 30101 MW; 684050E030488065 CRC64;
```

```
Query Match 20.7%; Score 6; DB 16; Length 295;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 INOLLC 9
Db 176 INOLLC 181

RESULT 27
O54221
ID O54221 PRELIMINARY; PRT; 300 AA.
AC O54221;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PROTEASE.
GN EGIP.
OS Staphylococcus epidermidis.
OG Plasmid pCH01.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1282;
[1]
SEQUENCE FROM N.A.
RC STRAIN=BN 280;
RX MEDLINE=98394975; PubMed=9726851;
RA Heldrich C., Pag U., Josten M., Metzger J., Jack R.W., Bierbaum G.,
RA Jung G., Sahl H.G.;
RT "Isolation, characterization and sequence of the novel lantibiotic
RT epicidin 280 and its biosynthetic gene cluster."
RL Appl. Environ. Microbiol. 64:3140-3146(1998).
DR EMBL; Y14023; CAA74349.1; -.
DR MEROPS; S08.0PA; -.
DR InterPro; IPR00209; Peptidase_S8.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PROSITE; PS00136; SUBTILASE_ASP; UNKNOWN_1.
KW Plasmid.
SQ SEQUENCE 300 AA; 34212 MW; D7DB1141A0A6AA66 CRC64;

Query Match 20.7%; Score 6; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 QLCKL 11
Db 220 QLCKL 225

RESULT 28
O06279
ID O06279 PRELIMINARY; PRT; 303 AA.
AC O06279;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 31.1 KDA PROTEIN.
GN RV3603C OR MTCY07H7B.19.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
[1]
SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
```

RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 DR EMBL: Z95557; CAB08941.1; -.
 DR Tuberculin; RV3603c; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 303 AA; 31104 MW; 3AB6E3B787E8A990 CRC64;

Query Match 20.7%; Score 6; DB 16; Length 303;
 Best Local Similarity 100.0%; Pred.No. 77;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDSELA 18
 |||||
 Db 78 TDSELA 83

RESULT 29
 O35100 PRELIMINARY; PRT; 308 AA.
 AC O35100;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PROTOCADHERIN 6 (FRAGMENT).
 OS Rattus rattus (Black rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 ON NCBI_TaxID=10117;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Obata S., Sago H., Nozomu M., Davidson M., St John T., Suzuki S.T.;
 RT "A common protocadherin tail: multiple protocadherins share the same
 RT regions in their cytoplasmic domains and are expressed in different
 RT segments of brain.";
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 2 CADHERIN DOMAINS.
 DR EMBL: AB005802; BAA22078.1; -.
 DR InterPro: IPR002126; Cadherin.
 DR Pfam: PF00028; cadherin; 1.
 DR PRINTS: PR00205; CADHERIN.
 DR SMART: SM00112; CA: 1.
 DR PROSITE: PS00232; CADHERIN_1; UNKNOWN_1.
 DR PROSITE: PS02368; CADHERIN_2; 2. Glycoprotein.
 KW Calcium-binding; Cell adhesion; Glycoprotein.
 FT NON_TER 1
 FT NON_TER 308 308
 SQ SEQUENCE 308 AA; 33343 MW; BB9EA7AEFE18AEAO CRC64;

Query Match 20.7%; Score 6; DB 11; Length 308;
 Best Local Similarity 100.0%; Pred.No. 78;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELAPRS 21
 |||||
 Db 57 ELAPRS 62

RESULT 30
 Q91TN5 PRELIMINARY; PRT; 313 AA.
 AC Q91TN5;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE T53.
 OS Tupaia herpesvirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Betaherpesvirinae.
 OX NCBI_TaxID=10397;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2;
 RX MEDLINE=21211637; PubMed=11312357;
 RA Bahr U., Darai G.;
 RT "Analysis and Characterization of the Complete Genome of Tupaia (Tree
 RT Shrew) Herpesvirus.";
 RL J. Virol. 75:4854-4870(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2;
 RA Darai G., Bahr U.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF281817; AAK57102.1; -.
 SQ SEQUENCE 313 AA; 35213 MW; BEAEBB9D275FBF07 CRC64;

Query Match 20.7%; Score 6; DB 12; Length 313;
 Best Local Similarity 100.0%; Pred.No. 80;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CINQLL 8
 |||||
 Db 242 CINQLL 247

RESULT 31
 P95153 PRELIMINARY; PRT; 346 AA.
 ID P95153;
 AC P95153;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ADHA.
 GN ADHA OR RV1862 OR MTCV359.11.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Teklaia F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 CC -1- COPACTOR: REQUIRES ZINC FOR ITS ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
 CC FAMILY.
 DR EMBL: Z83859; CAB06125.1; -.
 DR Tuberculin; RV1862; -.
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR Pfam: PF00107; adh_zinc; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 KW Complete proteome; Oxidoreductase; Zinc.
 SQ SEQUENCE 346 AA; 36631 MW; 107AD0EF4256812C CRC64;

Query Match 20.7%; Score 6; DB 16; Length 346;
 Best Local Similarity 100.0%; Pred.No. 87;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

QY 14 DSELAP 19
Db 157 DSELAP 162

RESULT 32
Q9LRX6 PRELIMINARY; PRT; 359 AA.
AC Q9LRX6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE SIMILARITY TO HETI PROTEIN FROM ANABAENA PCC7120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty p1 and TAC
RT clones."
RL DNA Res. 7:131-135(2000).
DR EMBL: AB028609; BAB02895.1; -.
SQ SEQUENCE 359 AA; 40668 MW; C6042D199BBEF3FD CRC64;

Query Match 20.7%; Score 6; DB 10; Length 359;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 LPTDSE 16
Db 89 LPTDSE 94

RESULT 33
Q9RZB3 PRELIMINARY; PRT; 364 AA.
AC Q9RZB3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE THYMIDINE DIPHOSPHOGLUCOSE 4,6-DEHYDRATASE.
GN DRA0041.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1."
RL Science 286:1571-1577(1999).
DR EMBL: AE001862; AAF12268.1; -.
DR HSSP: P27830; 1BKX.

```

```

DR TIGR: DRA0041; -.
DR InterPro: IPR001064; Crystallin.
DR InterPro: IPR001509; Epimerase.
DR Pfam: PF01370; Epimerase; 1.
KW PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 364 AA; 40885 MW; A3C2F6311AF469EC CRC64;

Query Match 20.7%; Score 6; DB 16; Length 364;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELAPRS 21
Db 155 ELAPRS 160

RESULT 34
Q9N0E8 PRELIMINARY; PRT; 391 AA.
AC Q9N0E8;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE P47-PHOX.
GN P47PHOX.
OS Tursiops truncatus (Atlantic bottle-nosed dolphin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;
OC Tursiops.
OX NCBI_TaxID=9739;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21109079; PubMed=11182145;
RA Inoue Y., Itou T., Jimbo T., Sakai T., Ueda K., Imajoh-Ohmi S.;
RT "Molecular cloning and identification of bottle-nosed dolphin
RT p40(phox), p47(phox) and p67(phox).";
RL Vet. Immunol. Immunopathol. 78:21-33(2001).
DR EMBL: AE035594; BAA96544.1; -.
DR HSSP: P14598; IGD5.
DR InterPro: IPR001655; P47PHOX.
DR InterPro: IPR001683; PX.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00018; SH3; 2.
DR PRINTS: PR00498; P47PHOX.
DR PRINTS: PR00452; SH3DOMAIN.
DR SMART: SM00312; PX; 1.
DR SMART: SM00326; SH3; 2.
DR PROSITE: PS00002; SH3; 2.
SQ SEQUENCE 391 AA; 44766 MW; F09E96D6B74B7C7C CRC64;

Query Match 20.7%; Score 6; DB 6; Length 391;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KLPTDS 15
Db 126 KLPTDS 131

RESULT 35
Q95L71 PRELIMINARY; PRT; 392 AA.
AC Q95L71;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NADPH OXIDASE CYTOSOLIC PROTEIN P47PHOX.
OS Bison bison (American bison).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

```

OC Bovidae; Bovinae; Bison.
OX NCBI_TaxID=9901;

RN [1]
RP SEQUENCE FROM N.A.
RA Gauss K.A., Bunger P.L., Quinn M.T.;
RT "Cloning of NADPH oxidase genes from North American bison."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF411138; AAL1188.1; -.
SQ SEQUENCE 392 AA; 45202 MW; E6CF3F360E7EDB44 CRC64;

Query Match 20.7%; Score 6; DB 6; Length 392;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KLPTDS 15
Db 126 KLPTDS 131

RESULT 36

ID O06038 PRELIMINARY; PRT; 396 AA.
AC O06038
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE EPSJ.
GN EPSJ.
OS Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).

OG Plasmid pNZ4000.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
OX NCBI_TaxID=1359;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIZO B40;
RX MEDLINE=97303093; PubMed=9159524;
RA van Kranenburg R., Marugg J.D., van Swam I.I., Willem N.J.,
RA de Vos W.M.;
RT "Molecular characterization of the plasmid-encoded eps gene cluster
essential for exopolysaccharide biosynthesis in Lactococcus lactis.";
RL Mol. Microbiol. 24:387-397(1997).

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NIZO B40;
RX MEDLINE=99084975; PubMed=9864348;
RA van Kranenburg R., van Swam I.I., Marugg J.D., Kleerebezem M.,
RA de Vos W.M.;
RT "Exopolysaccharide biosynthesis in Lactococcus lactis NIZO B40:
functional analysis of the glycosyltransferase genes involved in
synthesis of the polysaccharide backbone.";
RL J. Bacteriol. 181:338-340(1999).

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=NIZO B40;
RX MEDLINE=20153812; PubMed=10686131;
RA van Kranenburg R., Kleerebezem M., de Vos W.M.;
RT "Nucleotide sequence analysis of the lactococcal EPS plasmid
pNZ4000.";
RL Plasmid 43:130-136(2000).
DR EMBL; AF036485; AAC45237.1; -.
KW Plasmid.
SQ SEQUENCE 396 AA; 47514 MW; 350BF5236B9148B6 CRC64;

Query Match 20.7%; Score 6; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 LPTDSE 16
Db 212 LPTDSE 217

RESULT 37

ID Q9X5R5 PRELIMINARY; PRT; 404 AA.
AC Q9X5R5;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MITG.
GN MITG.
OS Streptomyces lavendulae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 2564;
RX MEDLINE=99201491; PubMed=10099135;
RA Mao Y., Varoglu M., Sherman D.H.;
RT "Molecular characterization and analysis of the biosynthetic gene
cluster for the antitumor antibiotic mitomycin C from Streptomyces
lavendulae NRRL 2564.";
RL Chem. Biol. 6:251-263(1999).
DR EMBL; AF127374; AAD28465.1; -.
DR InterPro; IPR000683; GFO_IDH_Moca.
DR InterPro; IPR004104; GFO_IDH_Moca_C.
DR Pfam; PF01408; GFO_IDH_Moca; 1.
DR Pfam; PF02894; GFO_IDH_Moca_C; 1.
DR PFam; PF02894; GFO_IDH_Moca_C; 1.
SQ SEQUENCE 404 AA; 43402 MW; E8FAE7A8899AB307 CRC64;

Query Match 20.7%; Score 6; DB 2; Length 404;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 ELAPRS 21
Db 67 ELAPRS 72

RESULT 38

ID Q9SG76 PRELIMINARY; PRT; 412 AA.
AC Q9SG76;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE GALACTOKINASE-LIKE PROPEIN.
GN T7M13.22.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC T7M13 genomic sequence.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC011708; AAF19579.1; -.
DR InterPro; IPR001459; Mv.gal.kin.
DR InterPro; IPR003880; Phosphopant_attach.
DR PRINTS; PR00959; MvGALKINASE.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
KW Kinase.
SQ SEQUENCE 412 AA; 44097 MW; 7D79A2D611D85D57 CRC64;

Query Match 20.7%; Score 6; DB 10; Length 412;

Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSEL 17
|||||

DB 4 PTDSEL 9

RESULT 39

Q54823 PRELIMINARY; PRT; 438 AA.
AC Q54823;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CYTOCHROME P450 131A1 (DNRP PROTEIN).
GN CYP131A1 OR DNRQ.
OS Streptomyces peucetius.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1950;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96062261; PubMed=7592454;
RA Otten S.L., Liu X., Ferguson J., Hutchinson C.R.;
RT "Cloning and characterization of the Streptomyces peucetius dnrQ
RT genes encoding a daunorubicin biosynthesis enzyme and a glycosyl
RT transferase involved in daunorubicin biosynthesis.";
RL J. Bacteriol. 177:6688-6692(1995).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; L47164; AAD15266.1; -.
DR InterPro; IPR001128; Cyt_P450.
DR Pfam; PF00067; p450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 438 AA; 46316 MW; A535F0B3611663B8 CRC64;

Query Match 20.7%; Score 6; DB 2; Length 438;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PTDSEL 17
|||||

DB 12 PTDSEL 17

RESULT 40

Q9AV62 PRELIMINARY; PRT; 441 AA.
AC Q9AV62;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PUTATIVE DISEASE RESISTANCE PROTEIN, 5'-PARTIAL (FRAGMENT).
GN OSJNBA0006L06.23.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Burr P.C.,
RA Hsiao J., Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,
RA Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,
RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 10 BAC OSJNba0006L06 genomic sequence.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC022457; AAK27812.2; -.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR; 12.

DR PRINTS; PRO0019; LEURICHRPT.
DR SMART; SM00369; LRR_typ; 7.
FT NON_TER 1
SQ SEQUENCE 441 AA; 48643 MW; ABF6366E769B6973 CRC64;

Query Match 20.7%; Score 6; DB 10; Length 441;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CINQLL 8
|||||

DB 427 CINQLL 432

RESULT 41

Q9Z599 PRELIMINARY; PRT; 442 AA.
ID Q9Z599;
AC Q9Z599;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE T4B21.10 PROTEIN.
GN T4B21.10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Abbott A., Kock J., Lehnert L.;
RT "The sequence of A. thaliana T4B21.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA WashU;
RT "The A. thaliana Genome Sequencing Project.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Waterston R.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AF118223; AAD03450.1; -.
DR InterPro; IPR003663; sub.transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PRO0171; SUGTRNSPORT.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane.
SQ SEQUENCE 442 AA; 48453 MW; 8A3FF353185A8659 CRC64;

Query Match 20.7%; Score 6; DB 10; Length 442;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SELAPR 20
|||||

DB 175 SELAPR 180

RESULT 42

Q9M100 PRELIMINARY; PRT; 461 AA.
ID Q9M100
AC Q9M100;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PUTATIVE SUGAR TRANSPORTER.
 GN AT4G04750.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,
 RA Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC EMBL; AL161501; CAB80840.1; -.
 DR InterPro: IPR003662; sub.transporter.
 DR InterPro: IPR003663; Sugar.transpster.
 DR PRINTS: PR00171; SUGRTRNSPRT.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Sugar transport; Transmembrane.
 SQ SEQUENCE 461 AA; 49470 MW; 652408EDD93DA3CA CRC64;

Query Match 20.7%; Score 6; DB 10; Length 461;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 SELAPR 20
 DB 131 SELAPR 136
 |||||

RESULT 43
 Q96TH4
 ID Q96TH4 PRELIMINARY; PRT; 499 AA.
 AC Q96TH4
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE TAKA-AMYLASE A (EC 3.2.1.1) (ALPHA-AMYLASE).
 GN AMYA OR AMY1.
 OS Aspergillus oryzae, and
 OS Aspergillus flavus.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5062, 5059;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=A.oryzae; STRAIN=RIB40;
 RA MEDLINE=20289310; PubMed=10830498;
 RA Gomi K., Akeno T., Minetoki T., Ozeki K., Kumagai C., Okazaki N.,
 RA Imura Y.;
 RT "Molecular cloning and characterization of a transcriptional activator
 RT gene, amyR, involved in the amylolytic gene expression in Aspergillus
 RT oryzae."
 RL Biosci. Biotechnol. Biochem. 64:816-827(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=A.flavus; STRAIN=86-10D;
 RA Fakhoury A.M., Woloshuk C.P.;
 RT "Amy1, the alpha-amylase gene of Aspergillus flavus: Involvement in
 RT aflatoxin biosynthesis in maize kernels."
 RL Phytopathology 89:908-914(1999).
 DR EMBL; AB021876; BAA95703.1; -.
 DR EMBL; AF139925; AAF14264.1; -.
 SQ SEQUENCE 499 AA; 54804 MW; EEF42ADA71D20DA9 CRC64;

Query Match 20.7%; Score 6; DB 3; Length 499;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 12 PTDSEL 17
 DB 374 PTDSEL 379
 |||||
 RESULT 44
 Q9VL48
 ID Q9VL48 PRELIMINARY; PRT; 503 AA.
 AC Q9VL48;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE CG13131 PROTEIN.
 GN CG13131.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Balwle R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J.D., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M.C., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhong G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003627; AAF52849.1; -.
 DR Flybase; FBgn0032175; CG13131.
 DR InterPro; IPR002965; P_rich_extensn.
 DR PRINTS; PR01217; PRICHEXTENS.
 SQ SEQUENCE 503 AA; 55716 MW; 74965AA4384B6F413 CRC64;

Query Match

20.7%; Score 6; DB 5; Length 503;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDSELA 18
|||||
Db 308 TDSELA 313

RESULT 45
Q93A34 PRELIMINARY; PRT; 504 AA.
AC Q93A34;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE OPUBB.
GN OPUBB.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RA Sleator R.D., Hill C.;
RT "Opus operon of Listeria monocytogenes LQ28.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF432069; AAL17617.1; --
SQ SEQUENCE 504 AA; 55682 MW; 50AA7893BD4E96BD CRC64;

Query Match 20.7%; Score 6; DB 2; Length 504;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDSELA 18
|||||
Db 420 TDSELA 425

RESULT 46
Q92BT3 PRELIMINARY; PRT; 504 AA.
AC Q92BT3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE LIN1461 PROTEIN.
GN LIN1461.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / SEROVAR 6A;
RX PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Blocker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetoui F., Couve E., de Barrois A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krest J., Kuhn M., Kunst F., Kurapkak G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjar H.,
RA Nordis G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL: AL596168; CAC96692.1; --
DR Listlist; LIN1461; --
KW Complete proteome.
SQ SEQUENCE 504 AA; 55763 MW; CBA6DDBE0E4D8E63 CRC64;

Query Match 20.7%; Score 6; DB 16; Length 504;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDSELA 18
|||||
Db 420 TDSELA 425

RESULT 47
Q98ALO PRELIMINARY; PRT; 521 AA.
AC Q98ALO;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE TRANSPOSASE.
GN MLL5961.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL: AP003007; BAB52319.1; --
DR InterPro: IPR002560; Transposase_12.
DR Pfam: PF01610; Transposase_12; 1.
KW Complete proteome.
SQ SEQUENCE 521 AA; 57850 MW; 027FDD40BAB02763 CRC64;

Query Match 20.7%; Score 6; DB 16; Length 521;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TDSELA 18
|||||
Db 469 TDSELA 474

RESULT 48
Q9VUJ3 PRELIMINARY; PRT; 602 AA.
AC Q9VUJ3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CG3020 PROTEIN.
GN CG3020.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foshier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Hostin D., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of *Drosophila melanogaster*.";
 RT Science 287:2185-2195(2000).
 RL EMBL; AE003532; AAF49683.1; -;
 DR HSSP; P21707; 1BYN.
 DR FlyBase; FBgn0036464; CG3020.
 DR InterPro; IPR0000008; C2.
 DR InterPro; IPR002149; LRI.
 DR Pfam; PF00168; C2; 2.
 DR PRINTS; PR00399; SYNAPTOTAGMIN.
 DR SMART; SM00239; C2; 2.
 DR PROSITE; PS00004; C2_DOMAIN_2; 2.
 SQ SEQUENCE 602 AA; 64957 MW; C73BCD777932BDB6 CRC64;

Query Match 20.7%; Score 6; DB 5; Length 602;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 11 LPTDSE 16
 Db 109 LPTDSE 114

RESULT 49
 Q22187 PRELIMINARY; PRT; 683 AA.
 AC Q22187;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DE T05A1.3 PROTEIN.
 GN T05A1.3.
 OS *Caenorhabditis elegans*.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; *Caenorhabditis*.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lloyd C.R.;

RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode *C.elegans*: A platform for
 RT investigating biology.";
 RL Science 282:2012-2018(1998).
 DR EMBL; Z68219; CAA92480.1; -;
 DR InterPro; IPR001611; LRR_out.
 DR Pfam; PF00560; LRR; 9.
 DR PRINTS; PR00019; LEURICHRPT.
 DR SMART; SM00370; LRR; 4.
 SQ SEQUENCE 683 AA; 77437 MW; 918E69A023421B43 CRC64;

Query Match 20.7%; Score 6; DB 5; Length 683;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 5 NOLLICK 10
 Db 367 NOLLICK 372

RESULT 50
 Q24213 PRELIMINARY; PRT; 704 AA.
 ID Q24213
 AC Q24213;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DE P21-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE P21-ACTIVATED PROTEIN KINASE 1 DPAK1.
 GN PAK OR DPAK1 OR CG10295.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; *Drosophila*.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chernoff J.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; U56080; AAB01209.1; -;
 DR HSSP; P24941; 1BUH.
 DR FlyBase; FBgn0014001; Pak.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR000095; PAK_box_P21_Rho_bindng.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR Pfam; PF00786; PBD; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR SMART; SM00285; PBD; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS50108; GBD; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 704 AA; 76189 MW; 12364B654429AA6F CRC64;

Query Match 20.7%; Score 6; DB 5; Length 704;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 12 PTDSEL 17
 Db 176 PTDSEL 181

Search completed: August 22, 2002, 11:20:34
Job time: 221 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: August 22, 2002, 11:19:39 ; Search time 13.5 seconds
(without alignments)
80.307 Million cell updates/sec

Title: US-09-856-199-2
Perfect score: 28
Sequence: 1 TGSATXQCKENLPCLCSXQGLRAEENI 28
Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0
Searched: 105224 seqs, 38719550 residues
Word size : 0
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Listing first 1200 summaries
Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	39.3	381	1 SELP_HUMAN	P49908 homo sapien
2	8	28.6	380	1 SELP_MOUSE	P70274 mus musculus
3	8	28.6	385	1 SELP_RAT	P25236 rattus norv
4	6	21.4	142	1 ATPE_PASMU	Q96kw0 pasteurella
5	6	21.4	221	1 YC24_GALSU	P35912 galdieria s
6	6	21.4	255	1 YHFD_BACSU	P39640 bacillus su
7	6	21.4	281	1 TRPA_CORGL	P08562 corynebacte
8	6	21.4	297	1 TFDA_BURSR	Q45423 burkholderi
9	6	21.4	297	1 YC24_ANTSP	Q02857 antithamnio
10	6	21.4	301	1 YGL4_BACST	P32814 bacillus st
11	6	21.4	360	1 UPP_TREPA	O83462 treponema p
12	6	21.4	394	1 SYVC_YEAST	P36421 saccharomyc
13	6	21.4	402	1 SELP_BOVIN	P49907 bos taurus
14	6	21.4	405	1 IF5_YEAST	P38431 saccharomyc
15	6	21.4	410	1 EFTU_CODFR	P50372 codium frag
16	6	21.4	476	1 ATPB_GALSU	Q08807 galdieria s
17	6	21.4	480	1 EXON_HSVJ7	P52346 human herpe
18	6	21.4	480	1 Y074_SYNV3	Q55790 synecocyst
19	6	21.4	483	1 YC24_CYACA	O9L1X2 cyanidium c
20	6	21.4	483	1 YC24_GUITH	O78473 guillardia
21	6	21.4	484	1 DF5L_HUMAN	P57764 homo sapien
22	6	21.4	486	1 YC24_CYAPA	P48260 cyanophora
23	6	21.4	487	1 GLRK_CHICK	P19439 gallus gall
24	6	21.4	487	1 YC24_PORPU	P51240 porphyra pu
25	6	21.4	495	1 SUFB_ECOLI	P77522 escherichia
26	6	21.4	500	1 PVCC_PSEAE	O30372 pseudomonas
27	6	21.4	553	1 MIS_RAT	P49000 rattus norv
28	6	21.4	555	1 MIS_MOUSE	P27106 mus musculu
29	6	21.4	575	1 MIS_BOVIN	Q03972 bos taurus
30	6	21.4	578	1 HLY4_AERSA	Q08677 aeromonas s
31	6	21.4	684	1 XYNA_CALSR	P40944 caldicellul
32	6	21.4	718	1 CDGT_BACLI	P14014 bacillus li
33	6	21.4	755	1 BGAL_RHIME	Q59750 rhizobium m

Q9upn9	homo sapien	1127	1	TF1G_HUMAN
Q24167	drosophila	1507	1	S1MA_DROME
P51112	fugu rubrip	3148	1	HD_FUGRU
Q31909	brassica ol	36	1	PSAI_BRAOL
P17227	pisum sativ	40	1	PSAI_PEA
Q9xhl1	bacillus ce	69	1	GRK_BACCE
P55797	rattus norv	88	1	APC4_RAT
P31261	notophthalm	90	1	HXA2_NOTVI
P76127	escherichia	92	1	YDDX_ECOLI
Q05772	brassica na	96	1	A9_BRANA
Q11172	mycobacteri	97	1	Y508_MYCTU
P01710	homo sapien	111	1	LV2G_HUMAN
P40076	saccharomyc	114	1	YEVI_YEAST
Q9h156	thermoplasm	121	1	RL7A_THEAC
P50507	oryctolagus	124	1	APC4_RABIT
P20348	drosophila	135	1	JANA_DROME
P04451	saccharomyc	137	1	RL23_YEAST
O42867	schizosacch	139	1	RL23_SCHPO
P49690	arabidopsis	140	1	RL23_ARATH
Q93140	brugia mala	140	1	RL23_BRUMA
P48158	caenorhabdi	140	1	RL23_CAEEL
P48159	drosophila	140	1	RL23_DROME
P23131	homo sapien	140	1	RL23_HUMAN
Q07760	nicotiana t	142	1	ATPE_HAEIN
P43718	haemophilus	147	1	Y648_METJA
Q58064	methanococc	148	1	PA21_HUMAN
P04054	homo sapien	149	1	YB07_METJA
Q58507	methanococc	150	1	SP0A_BACCE
P52930	baillus ce	153	1	UCRP_BOVIN
O02741	bos taurus	153	1	YEC8_YEAST
P39989	saccharomyc	158	1	RS7_LEPBI
P20819	leptospira	160	1	Y178_HAEIN
P43961	haemophilus	165	1	YSEA_STACA
P47995	staphylococ	171	1	KCCG_HUMAN
Q13555	homo sapien	174	1	IL10_TRIVU
O97798	trichosurus	178	1	INMA_CITFR
P05701	citrobacter	183	1	YAG2_SCHPO
Q09867	schizosacch	186	1	FI02_ECOLI
P08315	escherichia	186	1	FI04_ECOLI
O05781	escherichia	189	1	K112_SPIOL
P02398	spinacia ol	190	1	Y711_TREPA
O83709	treponema p	191	1	WAP_MACEU
Q9n018	macropus eu	193	1	PAPJ_ECOLI
P17543	escherichia	193	1	PRST_ECOLI
P42189	escherichia	196	1	UIDR_ECOLI
O59431	escherichia	199	1	HXA2_BRARE
O42365	brachydanio	199	1	HL18_SULSO
Q9ux88	sulfolobus	202	1	FTHC_MOUSE
Q9d110	mus musculu	202	1	R13A_CAEEL
Q27389	caenorhabdi	206	1	KTHY_NEIMA
Q9jve7	neisseria m	206	1	KTHV_NEIMB
Q9k0d9	neisseria m	208	1	ACPD_BACSU
Q35022	baillus su	208	1	NARP_HAEIN
P44845	haemophilus	209	1	SDHA_LACSK
Q48838	lactobacill	211	1	RL1_HALMA
P12738	haloarcula	211	1	UREF_MYCTU
P50050	mycobacteri	219	1	CD53_HUMAN
P19397	homo sapien	220	1	RADB_PYRKO
P95347	pyrococcus	221	1	YMI4_ARCFU
O28069	archaeoglob	223	1	CT14_MOUSE
P16410	homo sapien	223	1	CT14_PIG
P09793	mus musculu	229	1	YHCG_ECOLI
Q9mxx7	sus scrofa	233	1	PSMA_THEAC
P45426	escherichia	236	1	RR2_MAIZE
P25156	thermoplasm	236	1	RR2_ORYSA
P16037	zea mays (m	236	1	RR2_WHEAT
P12145	oryza sativ	236	1	NEUM_HUMAN
P17933	tritium ae	238	1	YTEA_BACSU
P17677	homo sapien	239	1	UDP_ECOLI
P42408	baillus su	252	1	UDP_SALTY
P12758	escherichia	252	1	UDP_KLEAE
O33808	salmonella	253	1	
O08444	klebsiella	253	1	

545	4	14.3	96	1	CH10_HAEDU	P31296	haemophilus	618	112	1	VNFG_AZOV1	P16857	azotobacter
546	4	14.3	96	1	CH10_HAEIN	P43734	haemophilus	619	113	1	HV27_MOUSE	P01796	mus musculus
547	4	14.3	96	1	CH10_PASMU	O59686	pasteurella	620	113	1	HV28_MOUSE	P01797	mus musculus
548	4	14.3	96	1	CTC2_ACILW	O33951	acinetobact	621	113	1	HV29_MOUSE	P01798	mus musculus
549	4	14.3	96	1	ES98_DROME	P07420	drosophila	622	113	1	HV30_MOUSE	P01799	mus musculus
550	4	14.3	96	1	NRAM_TADGE	P03480	influenza a	623	113	1	HV31_MOUSE	P01800	mus musculus
551	4	14.3	96	1	VG05_HSVSA	P24912	herpesvirus	624	113	1	HV34_MOUSE	P01803	mus musculus
552	4	14.3	96	1	VGE_BPG4	P03640	bacterioph	625	113	1	RL12_HALVO	P41197	halobacteri
553	4	14.3	96	1	YL34_ARCFU	O28146	archaeoglob	626	113	1	RP5M_ACIGB	P33987	acinetobact
554	4	14.3	97	1	CH10_ECOLI	P05380	escherichia	627	113	1	YG35_BPWV4	O04775	lactococcus
555	4	14.3	97	1	CH10_XANMA	P95801	xanthomonas	628	113	1	YKGI_YEAST	P35727	saccharomyc
556	4	14.3	97	1	GLHA_RANCA	P80051	r glycoprot	629	114	1	HV01_CANFA	P01784	canis famil
557	4	14.3	97	1	PSAE_CHLRE	P12352	chlamydomon	630	114	1	HV3B_HUMAN	P01763	homo sapien
558	4	14.3	98	1	FERR_ANASP	P11053	anabaena sp	631	114	1	VE7_HP41	P27556	human papil
559	4	14.3	98	1	FERR_ANAVA	P46046	anabaena va	632	114	1	YG11_STRCO	O9fc38	streptomyce
560	4	14.3	98	1	HUPF_RHILV	P27651	rhizobium l	633	115	1	CCRN_HUMAN	P06307	homo sapien
561	4	14.3	98	1	PYI_RAT	P10631	rattus norv	634	115	1	CCRN_MACFA	P23662	macaca fasc
562	4	14.3	98	1	VE7_HPV65	Q07859	human papil	635	115	1	CCRN_MOUSE	P09240	mus musculus
563	4	14.3	99	1	APC3_MOUSE	P33622	mus musculus	636	115	1	CCRN_RAT	P01355	rattus norv
564	4	14.3	99	1	DEFC_AEDAE	P81603	aedes aegypt	637	115	1	FIXT_RHIME	O86464	rhizobium m
565	4	14.3	99	1	HVYA_METVO	Q03576	methanococc	638	115	1	HV32_MOUSE	P01801	mus musculus
566	4	14.3	99	1	SVS7_MOUSE	Q09098	mus musculus	639	115	1	HV3D_HUMAN	P01765	homo sapien
567	4	14.3	99	1	Y754_AQUAE	O66956	aquifex ae	640	115	1	HV3F_HUMAN	P01767	homo sapien
568	4	14.3	100	1	PAHO_MOUSE	P10601	mus musculus	641	115	1	IMMU_ECOLI	P08702	escherichia
569	4	14.3	100	1	VNB_INBHK	P16192	influenza b	642	115	1	RL19_PASMU	O9cle0	pasteurella
570	4	14.3	101	1	YHBQ_ECOLI	P45472	escherichia	643	115	1	Y115_ADE02	P03290	human adeno
571	4	14.3	101	1	PCHB_PSEAE	O51507	pseudomonas	644	115	1	YVER_VACC	P20558	vaccinia vi
572	4	14.3	101	1	PRGJ_SALTY	P41785	salmonella	645	116	1	HV05_CARAU	P19181	carassius a
573	4	14.3	101	1	VG7_BPB03	Q37887	bacterioph	646	116	1	HV3Q_HUMAN	P01778	homo sapien
574	4	14.3	101	1	Y346_METJA	O57792	methanococc	647	116	1	NLTF_CICAR	O23758	cicer ariet
575	4	14.3	101	1	YQJD_ECOLI	P42617	escherichia	648	116	1	NU3M_SCYCA	O79408	scyllorhinu
576	4	14.3	102	1	ELAC_TRIVU	Q29143	trichosurus	649	116	1	NU3M_SQUAC	O92247	squalus aca
577	4	14.3	103	1	MGP_CHICK	O42413	gallus gall	650	116	1	RS17_PYRAB	O9vlu5	pyrococcus
578	4	14.3	104	1	PTLA_STRMU	P26426	streptococc	651	116	1	RS17_PYRBO	O59426	pyrococcus
579	4	14.3	104	1	SWT3_ARATH	P55852	arabidopsi	652	116	1	YA72_METJA	O58472	methanococc
580	4	14.3	104	1	SR19_ARCFU	O29010	archaeoglob	653	117	1	DHSD_ARCFU	P22973	archaeoglob
581	4	14.3	104	1	VSH7_DICDI	P14327	dictyosteli	654	117	1	E313_ADECG	P22228	canine aden
582	4	14.3	104	1	Y019_NYVOP	O10279	orgyia pseu	655	117	1	GALA_COTJA	O9gm9	cholurnix co
583	4	14.3	105	1	ATPL_NYVOP	O59550	mycoplasma	656	117	1	GC5H_CHLUM	O9pkb2	chlamydia m
584	4	14.3	105	1	LAC5_MOUSE	P20764	mus musculus	657	117	1	GC5H_CHLTR	O84284	chlamydia t
585	4	14.3	105	1	LAC5_MUSSH	P20765	mus spretus	658	117	1	HV3C_HUMAN	P01764	homo sapien
586	4	14.3	105	1	RL12_SULAC	P08055	sulfolobus	659	117	1	HV62_MOUSE	P18533	mus musculus
587	4	14.3	105	1	RR10_PORPU	P51286	porphyra pu	660	117	1	NTPF_ENTHR	P43437	enterococcu
588	4	14.3	105	1	YMQ3_CAEEL	P34494	caenorhabdi	661	117	1	SMT3_SCHPO	O13351	schizosacch
589	4	14.3	106	1	TBCA_YEAST	P48606	saccharomyc	662	117	1	YBAA_ECOLI	P09161	escherichia
590	4	14.3	107	1	A9_ARATH	Q00762	arabidopsi	663	117	1	YFB9_YEAST	P43576	saccharomyc
591	4	14.3	107	1	REV_HV2ST	P20870	human immun	664	117	1	YIAB_ECOLI	P11286	escherichia
592	4	14.3	107	1	Y206_LISMO	P33383	listeria mo	665	117	1	YJ9K_YEAST	P47174	saccharomyc
593	4	14.3	107	1	Y423_RICPR	O92da9	rickettsia	666	118	1	V118_ASFB7	P18556	african swi
594	4	14.3	107	1	YGP2_YEAST	P53103	saccharomyc	667	118	1	VPM_HAEIN	P44234	haemophilus
595	4	14.3	107	1	YI08_XYLFA	O9pob2	xyella fas	668	118	1	YG35_BPLLH	O04769	lactococcus
596	4	14.3	108	1	YF24_METJA	Q58919	methanococc	669	118	1	YI5C_ECOLI	P19770	escherichia
597	4	14.3	109	1	JHEA_TRINI	P30809	trichoplusi	670	119	1	EYAL_CHICK	O9yha0	gallus gall
598	4	14.3	109	1	VNST_BUNGE	P16992	bunyavirus	671	119	1	EYAL_CHICK	O9yha98	gallus gall
599	4	14.3	109	1	Y205_PASMU	P57825	pasteurella	672	119	1	EYA4_FUGUR	O9yhal	fugu rubrip
600	4	14.3	109	1	Y742_HAEIN	P44047	haemophilus	673	119	1	PMRA_LACLC	P27173	lactococcus
601	4	14.3	109	1	YBAB_HAEIN	P44711	haemophilus	674	119	1	VTU3_DROME	O06521	drosophila
602	4	14.3	110	1	NOS2_CAPHI	Q28314	capra hircu	675	119	1	YRAN_HAEIN	P45300	haemophilus
603	4	14.3	110	1	NBSP_HELVS	Q00572	helenium vi	676	120	1	HV3E_HUMAN	P01766	homo sapien
604	4	14.3	110	1	YHBJ_ACTAC	P96769	actinobacil	677	120	1	RL22_BORBU	P94272	borrelia bu
605	4	14.3	111	1	GDIR_CAVPO	P80237	cavia porce	678	120	1	RNPA_MYCLE	P46610	mycobacteri
606	4	14.3	111	1	HV35_MOUSE	P01804	mus musculus	679	120	1	VP3_CAV26	P54095	chicken ane
607	4	14.3	111	1	IBP5_BOVIN	Q05717	bos taurus	680	120	1	YACL_ECOLI	P45567	escherichia
608	4	14.3	111	1	LECS_RANJA	P18839	rana japoni	681	120	1	YGXO_YEAST	P53082	saccharomyc
609	4	14.3	111	1	LV2D_HUMAN	P01707	homo sapien	682	121	1	KDGL_ECOLI	P00556	escherichia
610	4	14.3	111	1	OL7H_MOUSE	Q60893	mus musculus	683	121	1	RA14_CANAL	P53709	candida alb
611	4	14.3	111	1	Y847_TREPA	O83819	treponema p	684	121	1	YI24_MYCTU	O50609	mycobacteri
612	4	14.3	111	1	YCV0_YEAST	P25638	saccharomyc	685	122	1	C556_AGRTB	P00139	agrobacteri
613	4	14.3	112	1	PEDB_PEDAC	P36496	pediococcus	686	122	1	C556_AGRTC	P00140	agrobacteri
614	4	14.3	112	1	PSAN_WAIZE	O65107	zea mays {m	687	122	1	HV20_MOUSE	P01789	mus musculus
615	4	14.3	112	1	RLA4_CLAHE	P42039	cladosporiu	688	122	1	HV21_MOUSE	P01790	mus musculus
616	4	14.3	112	1	VNFG_AZOCB	P15333	azotobacter	689	122	1	HV3G_HUMAN	P01768	homo sapien
617	4	14.3	112	1	VNFG_AZOSA	O68951	azotobacter	690	122	1	HV3H_HUMAN	P01769	homo sapien

				ALIGNMENTS			
				RESULT	1	SELP_HUMAN	STANDARD; PRT; 381 AA.
1129	4	14.3	180	Q9xd17	leptospira	SELP_HUMAN	
1130	4	14.3	180	Q06691	autographa	ID	
1131	4	14.3	180	P44245	haemophilus	AC	
1132	4	14.3	181	Q91611	salmonella	DT	
1133	4	14.3	181	Q44477	caenorhabdi	DT	
1134	4	14.3	181	Q56221	thermus aqu	DT	
1135	4	14.3	181	Q51122	arabidopsis	DT	
1136	4	14.3	181	O51122	borrelia bu	DE	
1137	4	14.3	181	O83696	treponema p	GN	
1138	4	14.3	181	P34565	caenorhabdi	OS	
1139	4	14.3	182	P46103	plasmodium	OC	
1140	4	14.3	182	P53521	proteus mir	OC	
1141	4	14.3	182	O69980	streptomyce	OX	
1142	4	14.3	183	P41866	zea mays (m	NCBI_TaxID=9606;	
1143	4	14.3	183	P37380	podocoryne	[1]	
1144	4	14.3	183	Q9ayp4	oryza sativ	SEQUENCE FROM N.A.	
1145	4	14.3	183	P41366	gallus gall	TISSUE=Liver, and Heart;	
1146	4	14.3	183	P76270	escherichia	MEDLINE=931133823; PubMed=8421687;	
1147	4	14.3	184	O84372	chlamydia t	Hill K.E., Lloyd R.S., Burk R.F.;	
1148	4	14.3	184	P76495	escherichia	"Conserved nucleotide sequences in the open reading frame and 3'	
1149	4	14.3	185	P35318	homo sapien	untranslated region of selenoprotein P mRNA.";	
1150	4	14.3	185	Q43359	bordeletia	Proc. Natl. Acad. Sci. U.S.A. 90:537-541(1993).	
1151	4	14.3	185	P54152	fragaria an	[2]	
1152	4	14.3	185	P40022	saccharomyc	PARTIAL SEQUENCE.	
1153	4	14.3	185	P38812	saccharomyc	MEDLINE=94191007; PubMed=8142465;	
1154	4	14.3	185	Q12425	saccharomyc	Aakesson B., Bellow T., Burk R.F.;	
1155	4	14.3	186	P14123	haloarcula	"Purification of selenoprotein P from human plasma.";	
1156	4	14.3	186	P35977	measles vir	Biochim. Biophys. Acta 1204:243-249(1994).	
1157	4	14.3	186	P03424	measles vir	[3]	
1158	4	14.3	186	P26035	measles vir	CHARACTERIZATION.	
1159	4	14.3	186	Q00794	measles vir	MEDLINE=98413836; PubMed=9735174;	
1160	4	14.3	186	P37547	bacillus su	Mostert V., Lombeck I., Abel J.;	
1161	4	14.3	186	P17962	trypanosoma	"A novel method for the purification of selenoprotein P from human	
1162	4	14.3	186	P44214	haemophilus	plasma.";	
1163	4	14.3	187	P28609	borrelia bu	Arch. Biochem. Biophys. 357:326-330(1998).	
1164	4	14.3	187	O29779	archaeoglob	[4]	
1165	4	14.3	187	O55003	mus musculus	CHARACTERIZATION.	
1166	4	14.3	187	Q02762	rhodobacter	MEDLINE=20239644; PubMed=10775431;	
1167	4	14.3	187	P44905	haemophilus	Mostert V.;	
1168	4	14.3	188	O62827	bos taurus	"Selenoprotein P: properties, functions, and regulation.";	
1169	4	14.3	188	P53366	sus scrofa	Arch. Biochem. Biophys. 376:433-438(2000).	
1170	4	14.3	188	P28893	ictaluriid h	[5]	
1171	4	14.3	188	P06911	rattus norv	REVIEW.	
1172	4	14.3	188	Q07006	streptomyce	MEDLINE=95017128; PubMed=7931697;	
1173	4	14.3	188	P57281	buchnera ap	Burk R.F., Hill K.E.;	
1174	4	14.3	188	Q9cb42	mycobacteri	"Selenoprotein P. A selenium-rich extracellular glycoprotein.";	
1175	4	14.3	188	Q57066	haemophilus	J. Nutr. 124:1891-1897(1994).	
1176	4	14.3	188	P23217	staphylococ	"FUNCTION: MIGHT BE RESPONSIBLE FOR SOME OF THE EXTRACELLULAR	
1177	4	14.3	189	Q41001	pisum sativ	ANTIOXIDANT DEFENSE PROPERTIES OF SELENIUM OR MIGHT BE INVOLVED IN	
1178	4	14.3	189	P03601	cowpea chlo	THE TRANSPORT OF SELENIUM.	
1179	4	14.3	189	P87163	cardia alb	CC	
1180	4	14.3	189	O87883	mycobacteri	CC	
1181	4	14.3	190	Q9cb17	mycobacteri	CC	
1182	4	14.3	190	Q07247	mycobacteri	CC	
1183	4	14.3	190	Q99yc8	streptococc	CC	
1184	4	14.3	190	P72657	synecocyst	CC	
1185	4	14.3	190	P17959	trypanosoma	CC	
1186	4	14.3	190	P35925	streptomyce	CC	
1187	4	14.3	191	O50917	borrelia bu	CC	
1188	4	14.3	191	Q92118	helicobacte	CC	
1189	4	14.3	191	P56104	helicobacte	CC	
1190	4	14.3	191	Q9zje5	helicobacte	CC	
1191	4	14.3	191	P57800	pasteurella	CC	
1192	4	14.3	192	P19241	staphylococ	CC	
1193	4	14.3	192	P44783	haemophilus	CC	
1194	4	14.3	192	Q9phm8	campylobact	CC	
1195	4	14.3	192	P09a78	xyliella fas	CC	
1196	4	14.3	192	P04388	drosofila	CC	
1197	4	14.3	192	P06535	bacillus su	CC	
1198	4	14.3	192	O67344	aquifex aeo	CC	
1199	4	14.3	193	P12115	cyprinus ca	CC	
1200	4	14.3	193	P29234	mustela vis	CC	

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 CC EMBL; Z11793; CAA77836.1; -
 CC MIM; 601484; -
 CC Glycoprotein; Signal; Selenium; Selenocysteine; Plasma.
 FT SIGNAL 1 19
 FT CHAIN 20 381 SELENOPROTEIN P.
 FT DOMAIN 244 249 POLY-HIS.


```
RESULT 6
YWFD_BACSU
ID YWFD_BACSU STANDARD; PRT; 255 AA.
AC P39640;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical oxidoreductase ywfd (EC 1.-.-.-).
GN YWFD OR IPA-82D.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=95020537; PubMed=7934828;
RA Glaser P., Kunst F., Arnaud M., Coudart M.P., Gonzales W.,
RA Hullo M.F., Ionescu M., Lubochinsky B., Marcelino L., Moszer I.,
RA Presecan E., Santana M., Schneider E., Schweizer J., Vertes A.,
RA Rapoport G., Danchin A.;
RT "Bacillus subtilis genome project: cloning and sequencing of the 97
RT kb region from 325 degrees to 333 degrees.";
RL Mol. Microbiol. 10:371-384(1993).
CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X73124; CRA51638.1; -
DR EMBL; Z99123; CAB5799.1; -
DR PIR; S39737; S39737.
DR HSSP; P50162; 1AE1.
DR Subtilist; BG10628; ywfd.
DR InterPro; IPR002198; ADH_short.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; 1.
KW Hypothetical protein; Oxidoreductase; Complete proteome.
FT NP_BIND 11 33 NAD OR NADP (BY SIMILARITY).
FT ACT_SITE 154 154 BY SIMILARITY.
SQ SEQUENCE 255 AA; 27324 MW; 20AA2259BFB88C9B CRC64;

Query Match 21.4%; Score 6; DB 1; Length 255;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
Db 242 TGSALT 247

RESULT 7
TRPA_CORGL
ID TRPA_CORGL STANDARD; PRT; 281 AA.
AC P06562;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tryptophan synthase alpha chain (EC 4.2.1.20).
GN TRPA.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;
OC Corynebacterium.
```

```
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87117512; PubMed=3808947;
RA Matsui K., Sano K., Ohtsubo E.;
RT "Complete nucleotide and deduced amino acid sequences of the
RT Brevibacterium lactofermentum tryptophan operon.";
RL Nucleic Acids Res. 14:10113-10114(1986).
RN [2]
RP SEQUENCE OF 161-281 FROM N.A.
RX MEDLINE=87277409; PubMed=3609747;
RA Sano K., Matsui K.;
RT "Structure and function of the trp operon control regions of
RT Brevibacterium lactofermentum, a glutamic-acid-producing bacterium.";
RL Gene 53:191-200(1987).
CC -!- FUNCTION: THE ALPHA SUBUNIT IS RESPONSIBLE FOR THE ALDOL CLEAVAGE
CC OF INDOLEGLYCEROL PHOSPHATE TO INDOLE AND GLYCERALDEHYDE 3-
CC PHOSPHATE.
CC -!- CATALYTIC ACTIVITY: L-serine + 1-(indol-3-yl)glycerol 3-phosphate
CC = L-tryptophan + glyceraldehyde 3-phosphate + H(2)O.
CC -!- PATHWAY: LAST (FIFTH) STEP IN BIOSYNTHESIS OF TRYPTOPHAN.
CC -!- SUBUNIT: TETRAMER OF TWO ALPHA AND TWO BETA CHAINS (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE TRPA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04960; CAA28628.1; -
DR EMBL; M16664; AAA83990.1; -
DR PIR; G24723; G24723.
DR PIR; C29458; C29458.
DR HSSP; P00929; 2WSY.
DR InterPro; IPR003009; FMN_enzyme.
DR InterPro; IPR002028; Trp_synthase_alpha.
DR Pfam; PF00290; trp_synth; 1.
DR ProDom; PD001535; TRP_synthase_alpha; 1.
DR PROSITE; PS00167; TRP_SYNTHASE_ALPHA; 1.
KW Tryptophan biosynthesis; Lyase.
SQ SEQUENCE 281 AA; 29222 MW; 99524638FF7C0FLE CRC64;

Query Match 21.4%; Score 6; DB 1; Length 281;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
Db 232 TGSALT 237

RESULT 8
TFDA_BURSR
ID TFDA_BURSR STANDARD; PRT; 297 AA.
AC Q45423;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Alpha-ketoglutarate-dependent 2,4-dichlorophenoxyacetate dioxygenase
DE (EC 1.14.-.-) (2,4-D dioxygenase).
GN TFDA.
OS Burkholderia sp. (strain RASC).
OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC Burkholderia.
OX NCBI_TaxID=69003;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96271695; PubMed=8779585;
```

RA Suwa Y., Wright A.D., Fukimori F., Nummy K.A., Hausinger R.P.,
RA Holben W.E., Forney L.J.,
RT "Characterization of a chromosomally encoded
RT 2,4-dichlorophenoxyacetic acid/alpha-ketoglutarate dioxygenase from
RT Burkholderia sp. strain RASC.",
RL Appl. Environ. Microbiol. 62:2464-2469(1996).
CC -1- FUNCTION: INVOLVED IN DEGRADATION OF 2,4-D, 2-METHYL-4-CHLOROPHEN-
CC OXYACETIC ACID AND 3-CHLOROBENZOIC ACID.
CC -1- COPACTOR: IRON AND ASCORBATE (BY SIMILARITY).
CC -1- PATHWAY: 3-CHLOROCATECHOL DEGRADATION.
CC -1- SIMILARITY: BELONGS TO THE TFDA FAMILY OF DIOXYGENASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U25717; AAB17363.1; -.
DR InterPro: IPR003819; Taud_Tfda.
DR Pfam; PF02668; Taud; 1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Dioxygenase;
KW Iron; Vitamin C.
SQ SEQUENCE 297 AA; 33638 MW; 17BF90A067C511FD CRC64;

Query Match 21.4%; Score 6; DB 1; Length 297;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
Db 162 QGLRAE 167

RESULT 9
YC24_ANTSP STANDARD; PRT; 297 AA.
ID YC24_ANTSP
AC Q02857;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein in ATPA 3'region (ORF 3) (Fragment).
GN YC24.
OS Antithamnion sp.
OG Chloroplast.
OC Eukaryota; Rhodophyta; Florideophyceae; Ceramiales; Ceramiaceae;
OC Antithamnion.
OX NCBI_TaxID=2767;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LB 95.79;
RX MEDLINE=93021132; PubMed=1404401;
RA Kostrewa M., Zetsche K.;
RT "Large ATP synthase operon of the red alga Antithamnion sp. resembles
RT the corresponding operon in cyanobacteria.",
RL J. Mol. Biol. 227:961-970(1992).
CC -1- SIMILARITY: BELONGS TO THE UPF0051 (YCF24) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X63382; CAA44986.1; -.
DR PIR; S37636; S37636.
DR Mendel; 6185; ANTsp:ycf24.1.
DR InterPro; IPR000825; UPF0051.

DR Pfam; PF01458; UPF0051; 1.
KW Chloroplast; Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 297 AA; 32894 MW; D9E5C4704042D763 CRC64;

Query Match 21.4%; Score 6; DB 1; Length 297;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
Db 123 TGSALT 128

RESULT 10
YGL4_BACST STANDARD; PRT; 301 AA.
ID YGL4_BACST
AC P32814;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 35.5 kDa protein in GLDA 3'region (ORF4).
OS Bacillus steaerothermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VAR. NON-DIASTATICUS;
RX MEDLINE=92184120; PubMed=1339360;
RA Mallinder P.R., Pritchard A., Moir A.;
RT "Cloning and characterization of a gene from Bacillus
RT steaerothermophilus var. non-diastaticus encoding a glycerol
RT dehydrogenase.",
RL Gene 110:9-16(1992).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M65289; AAA22479.1; -.
DR PIR; JS0681; JS0681.
KW Hypothetical protein.
SQ SEQUENCE 301 AA; 35522 MW; C0403C0A68A3CEBF CRC64;

Query Match 21.4%; Score 6; DB 1; Length 301;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GLRAEE 26
Db 218 GLRAEE 223

RESULT 11
UPP_TREPA STANDARD; PRT; 360 AA.
ID UPP_TREPA
AC O83462;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative uracil phosphoribosyltransferase (EC 2.4.2.9) (UMP
DE pyrophosphorylase) (UPPase).
GN UPP OR TP0448.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;

```

RN  SEQUENCE FROM N.A.
RC  STRAIN=NICHOLS;
RX  MEDLINE=98332770; PubMed=9665876;
RA  Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA  Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA  Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA  Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA  McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA  Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA  Venter J.C.;
RT  "Complete genome sequence of Treponema pallidum, the syphilis
RT  Spirochete.";
RL  Science 281:375-388(1998).
CC  -1- CATALYTIC ACTIVITY: UMP + diphosphate = uracil + 5-phospho-
CC  alpha-D-ribose 1-diphosphate.
CC  -1- PATHWAY: PYRIMIDINE SALVAGE PATHWAY.
CC  -1- SIMILARITY: BELONGS TO THE UPTASE FAMILY. VERY DIVERGENT.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AE001221; AAC65434.1; -
DR  TIGR; TP0448; -
DR  InterPro; IPR000836; Priboyltran.
DR  Pfam; PF00156; Priboyltran; 1.
KW  Hypothetical protein; Transferase; Glycosyltransferase;
KW  Complete proteome.
SQ  SEQUENCE 360 AA; 41017 MW; 9E2CAB2E2FA3768C CRC64;

Query Match          21.4%; Score 6; DB 1; Length 360;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  23 RAEANI 28
    |||||
DB  56 RAEANI 61

RESULT 12
SYCC YEAST
ID  SYCC YEAST          STANDARD;          PRT;          394 AA.
AC  P36421;
DT  01-JUN-1994 (Rel. 29, Created)
DT  01-JUN-1994 (Rel. 29, Last sequence update)
DT  01-NOV-1997 (Rel. 35, Last annotation update)
DE  Tyrosyl-tRNA synthetase, cytoplasmic (BC 6.1.1.1) (Tyrosyl-tRNA
DE  ligase) (TYRRS).
GN  TY51 OR MGM104 OR YGR185C OR G7522.
OS  Saccharomyces cerevisiae (Baker's yeast).
OC  Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC  Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX  NCBI_TaxID=4932;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=93286133; PubMed=8509419;
RA  Chow C.M., Rajbhandary U.L.;
RT  "Saccharomyces cerevisiae cytoplasmic tyrosyl-tRNA synthetase gene.
RT  Isolation by complementation of a mutant Escherichia coli suppressor
RT  tRNA defective in aminoacylation and sequence analysis.";
RL  J. Biol. Chem. 268:12855-12863(1993).
RN  [2]
RP  SEQUENCE FROM N.A.
RA  Guan M.-X., Chen X.-J., Clark-Walker G.D.;
RA  Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
RN  [3]
RP  SEQUENCE FROM N.A.

```

```

RC  STRAIN=S288C;
RX  MEDLINE=97279231; PubMed=9133739;
RA  Arroyo J., Garcia-Gonzalez M., Garcia-Saez M.I., Sanchez-Perez M.,
RA  Nombela C.;
RT  "DNA sequence analysis of a 23,002 bp DNA fragment of the right arm
RL  of Saccharomyces cerevisiae chromosome VII.";
RN  Yeast 13:357-363(1997).
RP  [4]
RP  SEQUENCE OF 1-36 FROM N.A.
RC  STRAIN=BJ926;
RX  MEDLINE=95087887; PubMed=7995524;
RA  Henry N.L., Campbell A.M., Feaver W.J., Poon D., Weil P.A.,
RA  Kornberg R.D.;
RT  "TFIIIF-TAF-RNA polymerase II connection.";
RL  Genes Dev. 8:2868-2878(1994).
CC  -1- CATALYTIC ACTIVITY: ATP + L-tyrosine + tRNA(Tyr) = AMP +
CC  diphosphate + L-tyrosyl-tRNA(Tyr).
CC  -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC  -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; LI2221; AAB59329.1; -
DR  EMBL; X71998; -; NOT_ANNOTATED_CDS.
DR  EMBL; Z72970; CAA97211.1; -
DR  EMBL; X99074; CAA67529.1; -
DR  EMBL; U13015; AAA61641.1; -
DR  PIR; A45999; A45999.
DR  SGD; S0003417; TYS1.
DR  InterPro; IPR002305; tRNA-synt_lb.
DR  InterPro; IPR001412; tRNA-synt_1.
DR  InterPro; IPR002307; tRNA-synt_tyr.
DR  Pfam; PF00579; tRNA-synt_lb; 1.
DR  PRINTS; PR1040; TRNASYNTHYR.
DR  PROSITE; PS00178; AA_TRNA_LIGASE_I; FALSE_NEG.
KW  Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT  SITE 48 56
FT  SITE 227 231 "HIGH" REGION.
SQ  SEQUENCE 394 AA; 44020 MW; 57E8DB9BE6D054B7 CRC64;

Query Match          21.4%; Score 6; DB 1; Length 394;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  11 ENLPSL 16
    |||||
DB  201 ENLPSL 206

RESULT 13
SELP_BOVIN
ID  SELP_BOVIN          STANDARD;          PRT;          402 AA.
AC  P49907; O19003;
DT  01-OCT-1996 (Rel. 34, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Selenoprotein P-like protein precursor.
GN  SEPP1 OR SELP.
OS  Bos taurus (Bovine).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC  Bovidae; Bovinae; Bos.
OX  NCBI_TaxID=9913;
RN  [1]
RP  SEQUENCE FROM N.A.
TISSUE=Cerebellum;

```

```

RX MEDLINE=95364621; PubMed=7637580;
RA Saitoh K., Saito N., Lee M.J., Fujii M., Kobayashi T., Sumino K.;
RT "Molecular cloning of cDNA encoding a bovine selenoprotein P-like
RT protein containing 12 selenocysteines and a (His-Pro) rich domain
RT insertion, and its regional expression.";
RL Brain Res. Mol. Brain Res. 30:301-311(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98019090; PubMed=9358058;
RA Fujii M., Saitoh K., Kobayashi T., Fujii S., Lee M.J., Sumino K.;
RT "Analysis of bovine selenoprotein P-like protein gene and availability
RT of metal responsive element (MRE) located in its promoter.";
RL Gene 199:211-217(1997).
CC -!- FUNCTION: IT CONSTITUTES A MAJOR SELENIUM POOL IN THE BRAIN AND
CC MAY PLAY AN IMPORTANT ROLE IN DEVELOPING AND/OR MODULATING THE
CC MORPHOLOGY OF NEURONS AND/OR GLIAL CELLS.
CC -!- SUBCELLULAR LOCATION: Extracellular (By similarity).
CC -!- TISSUE SPECIFICITY: BRAIN AND KIDNEY. MOST PROMINENTLY EXPRESSED
CC IN THE CEREBELLAR CORTEX, HIPPOCAMPUS AND OLFACTORY BULB.
CC -!- MISCELLANEOUS: THE SELENOCYSTEINES ARE ALL ENCODED BY THE OPAL
CC CODON, UGA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D25220; BAA04949.2; -.
DR EMBL; D88033; BAA23414.1; -.
DR EMBL; D88031; BAA23414.1; JOINED.
DR EMBL; D88032; BAA23414.1; JOINED.
DR HSPB; P04355; 4MT2.
KW Glycoprotein; Signal; Selenium; Selenocysteine; Repeat.
FT SIGNAL 1 19
FT CHAIN 20 402
FT SE_CYS 59 59
FT SE_CYS 297 297
FT SE_CYS 307 307
FT SE_CYS 338 338
FT SE_CYS 350 350
FT SE_CYS 363 363
FT SE_CYS 365 365
FT SE_CYS 372 372
FT SE_CYS 388 388
FT SE_CYS 390 390
FT SE_CYS 397 397
FT SE_CYS 399 399
FT SE_CYS 204 239
FT DOMAIN 260 266
FT POLY-HIS.
FT SRPQ -> KALE (IN REF. 2).
FT T -> P (IN REF. 2).
FT L -> V (IN REF. 2).
FT Y -> D (IN REF. 2).
SQ SEQUENCE 402 AA; 45018 MW; B7CF18751FBOE8FF CRC64;
H-P REPEATS.
Query Match 21.4%; Score 6; DB 1; Length 402;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 TGSALT 6
Db 332 TGSALT 337
RESULT 14
ID IF5_YEAST STANDARD; PRT; 405 AA.
AC P38431;
DT 01-OCT-1994 (Rel. 30, Created)

```

```

DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Eukaryotic translation initiation factor 5 (eIF-5).
GN TIF5 OR YPR041W OR YP3085.05
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-30; 326-340 AND 363-377.
RC STRAIN=X2180 / ATCC 26109;
RX MEDLINE=93252941; PubMed=8486705;
RA Chakravarti D., Maitra U.;
RT "Eukaryotic translation initiation factor 5 from Saccharomyces
RT cerevisiae. Cloning, characterization, and expression of the gene
RT encoding the 45-346-Da protein.";
RL J. Biol. Chem. 268:10524-10533(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=IL125-2B;
RX MEDLINE=98398050; PubMed=9730282;
RA Gaskiewicz-Staniorowska B., Skala J., Jasinski M., Grenson M.,
RA Goffeau A., Ulaszewski S.;
RT "Functional analysis of three adjacent open reading frames from the
RT right arm of yeast chromosome XVI.";
RL Yeast 14:1027-1039(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Badcock K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF GTP BOUND TO THE 40S
CC RIBOSOMAL INITIATION COMPLEX (40S MRNA.MET-TRNA[F], EIF-2.GTP) WITH
CC THE SUBSEQUENT JOINING OF A 60S RIBOSOMAL SUBUNIT RESULTING IN THE
CC RELEASE OF EIF-2 AND THE GUANINE NUCLEOTIDE. THE SUBSEQUENT
CC JOINING OF A 60S RIBOSOMAL SUBUNIT RESULTS IN THE FORMATION OF A
CC FUNCTIONAL 80S INITIATION COMPLEX (80S.MRNA.MET-TRNA[F]). EIF-5
CC IS ESSENTIAL FOR CELL VIABILITY.
CC -!- SUBUNIT: MONOMER.
CC -!- ALTERNATIVE PRODUCTS: TWO ISOFORMS MAY BE PRODUCED FROM THE USE OF
CC ALTERNATIVE INITIATION CODONS IN THE SAME READING FRAME.
CC -!- SIMILARITY: BELONGS TO THE EIF-2-BETA / EIF-5 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L10840; -. NOT_ANNOTATED_CDS.
DR EMBL; Z71255; CAA94989.1; -.
DR EMBL; Z68111; CAA92145.1; -.
DR EMBL; Z73616; CAA97991.1; -.
DR PIR; A46699; A46699.
DR SGD; S0006245; TIF5.
DR InterPro; IPR003307; eIF5C.
DR InterPro; IPR002735; eIF5_eIF2B.
DR Pfam; PF01873; eIF5_eIF2B; 1.
DR Pfam; PF02020; W2; 1.
DR ProDom; PD004078; eIF5_eIF2B; 1.
DR SMART; SM00515; eIF5C; 1.
KW Initiation factor; Protein biosynthesis; Alternative initiation;
KW GTP-binding.
FT CHAIN 1 405
FT EUKARYOTIC INITIATION FACTOR 5, LONG
FT ISOFORM.
FT EUKARYOTIC INITIATION FACTOR 5, SHORT
FT ISOFORM.
FT CHAIN 18 405
FT INIT_MET 18 18
FT DOMAIN 152 156
FT NP_BIND 27 34
FT POLY-LYS.
FT GTP (POTENTIAL).
FT ASP/GLU-RICH (ACIDIC).
FT DOMAIN 396 405

```

SQ SEQUENCE 405 AA; 45261 MW; 1A1DA563B4ED1B1F CRC64;

Query Match 21.4%; Score 6; DB 1; Length 405;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KENLPS 15
|||||

Db 263 KENLPS 268

RESULT 15

EFU_CODFR STANDARD; PRT; 410 AA.
AC P50372;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN TUFA.
OS Codium fragile.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Chlorophyta; Ulvophyceae; Caulerpaceae;
OC Codiaceae; Codium.
OX NCBI_TaxID=3133;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95392822; PubMed=7663757;
RA Delwiche C.F., Kuhsel M., Palmer J.D.;
RT "Phylogenetic analysis of tufa sequences indicates a cyanobacterial
origin of all plastids.";
RL Mol. Phylogenet. Evol. 4:110-128(1995).
CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
SYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U09427; AAA87687.1; -
CC HSSP; P02990; 1EFU.
CC Mendel; 13613; CODfr; TufA; 1.
CC InterPro; IPR000795; GTP_EFTU.
CC InterPro; IPR004161; GTP_EFTU_D2.
CC InterPro; IPR004160; GTP_EFTU_D3.
CC Pfam; PF00009; GTP_EFTU; 1.
CC Pfam; PF03144; GTP_EFTU_D2; 1.
CC Pfam; PF03143; GTP_EFTU_D3; 1.
CC PRINTS; PR00315; ELONGATNFT.
CC PROSITE; PS00301; EFATOR_GTP; 1.
CC Elongation factor; Protein biosynthesis; Chloroplast;
KW GTP-binding.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 81 85 GTP (BY SIMILARITY).
FT NP_BIND 136 139 GTP (BY SIMILARITY).
SQ SEQUENCE 410 AA; 45518 MW; 949A107D5519A56E CRC64;

Query Match 21.4%; Score 6; DB 1; Length 410;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
|||||

Db 172 TGSALT 177

RESULT 16

ATPB_GALSU STANDARD; PRT; 476 AA.
AC Q08807;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE ATP synthase beta chain (EC 3.6.3.14).
GN ATPB.
OS Galdieria sulphuraria.
OG Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;
OX Galdieria.
OX NCBI_TaxID=130081;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=14-1-1 / ISOLATE 107.79/GOETTINGEN;
RX MEDLINE=94033298; PubMed=8219057;
RA Kozrzewa M., Zetsche K.;
RT "Organization of plastid-encoded ATPase genes and flanking regions
including homologues of infB and tsf in the thermophilic red alga
Galdieria sulphuraria.";
RL Plant Mol. Biol. 23:67-76(1993).
CC -1- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
GRADIENT ACROSS THE MEMBRANE. THE BETA CHAIN IS THE CATALYTIC
SUBUNIT.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) -> ADP + phosphate +
H(+)(Out).
CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
HAS THREE MAIN SUBUNITS: A, B AND C.
CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC -1- SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X66698; CAA47241.1; -
CC PIR; S36412; S36412.
CC HSSP; P00829; 1BMF.
CC InterPro; IPR003593; AAA.
CC InterPro; IPR004100; ATP-synt_ab_N.
CC InterPro; IPR000793; ATPase_AB_C.
CC InterPro; IPR000194; ATPase_alpha_beta.
CC Pfam; PF00006; ATP-synt_ab; 1.
CC Pfam; PF00306; ATP-synt_ab_C; 1.
CC Pfam; PF02874; ATP-synt_ab_N; 1.
CC SMART; SM00382; AAA; 1.
CC PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
KW ATP synthesis; Chloroplast; Thylakoid; Membrane; CF(1);
FT NP_BIND 155 162 ATP (BY SIMILARITY).
SQ SEQUENCE 476 AA; 51966 MW; A30500E1062213AC CRC64;

Query Match 21.4%; Score 6; DB 1; Length 476;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6

Db 77 TGSALT 82


```
RESULT 17
EXON_HSV7J
ID EXON_HSV7J STANDARD; PRT; 480 AA.
AC P52346;
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alkaline exonuclease (EC 3.1.11.-).
GN U70.
OS Human herpesvirus (type 7 / strain J1) (HHV7).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=57278;
RN [1]
RP SEQUENCE FROM N.A.
RA Nicholas J.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES ALKALINE EXONUCLEASE
CC FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U43400; AAC54731.1; -.
DR InterPro; IPR001616; Herpes_alk_exo.
DR Pfam; PF01771; Herpes_alk_exo; 1.
DR PRINTS; PR00924; ALKEXNUCLASE.
KW Hydrolase; Nuclease; Exonuclease.
SQ SEQUENCE 480 AA; 55999 MW; CA3EAAAF48B5113E2 CRC64;

Query Match 21.4%; Score 6; DB 1; Length 480;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KENLPS 15
DB 384 KENLPS 389

RESULT 18
Y074_SYNY3
ID Y074_SYNY3 STANDARD; PRT; 480 AA.
AC Q55790;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 52.8 kDa protein SLR0074.
GN SLR0074.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE; 96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
CC -!- SIMILARITY: BELONGS TO THE UPF0051 (YCF24) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D64004; BAA10542.1; -.
DR InterPro; IPR000825; UPF0051.
DR Pfam; PF01458; UPF0051; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 480 AA; 52795 MW; 066845CB17A0F33A CRC64;

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D64004; BAA10542.1; -.
DR InterPro; IPR000825; UPF0051.
DR Pfam; PF01458; UPF0051; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 480 AA; 52795 MW; 066845CB17A0F33A CRC64;

Query Match 21.4%; Score 6; DB 1; Length 480;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSAT 6
DB 306 TGSAT 311

RESULT 19
YC24_GYACA
ID YC24_GYACA STANDARD; PRT; 483 AA.
AC Q9TLX2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 54.3 kDa protein YCF24.
GN YCF24 OR YCF32.
OS Cyanidium caldarium.
OG Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;
OC Cyanidium.
OX NCBI_TaxID=2771;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RK-1;
RX MEDLINE=20496959; PubMed=11040290;
RA Gloeckner G., Rosenthal A., Valentin K.;
RT "The structure and gene repertoire of an ancient red algal plastid
RT genome.";
RL J. Mol. Evol. 51:382-390(2000).
CC -!- SIMILARITY: BELONGS TO THE UPF0051 (YCF24) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF022186; AAF12948.1; -.
DR InterPro; IPR000825; UPF0051.
DR Pfam; PF01458; UPF0051; 1.
KW Chloroplast; Hypothetical protein.
SQ SEQUENCE 483 AA; 54151 MW; 8DE99BB7828D919B CRC64;

Query Match 21.4%; Score 6; DB 1; Length 483;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSAT 6
DB 309 TGSAT 314

RESULT 20
YC24_GUITH
ID YC24_GUITH STANDARD; PRT; 483 AA.
AC YC24_GUITH
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
```

```
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Hypothetical 54.0 kDa protein YCF24.
GN YCF24.
OS Guillardia theta (Cryptomonas phi).
OC Chloroplast.
OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
OX NCBI_TaxID=55529;
RN [1]
RP SEQUENCE FROM N.A.
RA Leitsch C.E.W., Kowallik K.V., Douglas S.E.;
RT "The atpA gene cluster of a cryptomonad, Guillardia theta: a piece in
RL the puzzle of chloroplast genome development.";
RN J. Phycol. 35:128-135(1999).
RP [2]
RP SEQUENCE FROM N.A.
RA Douglas S.E., Penny S.L.;
RT "The plastid genome of the cryptophyte alga, Guillardia theta:
RL complete sequence and conserved syntenic groups confirm its common
RT ancestry with red algae.";
RN J. Mol. Evol. 48:236-244(1999).
RP [1- SIMILARITY: BELONGS TO THE UPF0051 (YCF24) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF0411468; AAC35664.1; -
SQ SEQUENCE 483 AA; 54023 MW; 42C37B25F952BF76 CRC64;

Query Match 21.4%; Score 6; DB 1; Length 483;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGSALT 6
Db 309 TGSALT 314
|||||

RESULT 21
DF5L_HUMAN
ID DF5L_HUMAN STANDARD; PRT; 484 AA.
AC P57764;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE DFN5-like protein FLJ12150.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Magatsuma M., Hosoiri T., Kaku Y., Kodaaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Yamamoto S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamanoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC [1- SIMILARITY: BELONGS TO THE DFN5 FAMILY.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF0411468; AAC35664.1; -
SQ SEQUENCE 483 AA; 54023 MW; 42C37B25F952BF76 CRC64;

Query Match 21.4%; Score 6; DB 1; Length 483;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGSALT 6
Db 309 TGSALT 314
|||||

RESULT 22
YC24_CYAPA
ID YC24_CYAPA STANDARD; PRT; 486 AA.
AC P48260;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 54.3 kDa protein YCF24.
GN YCF24.
OS Cyanophora paradoxa.
OC Cyanelle.
OC Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.
OX NCBI_TaxID=2762;
RN [1]
RP SEQUENCE FROM N.A.
RA Stirewalt V.L., Michalowski C.B., Loeffelhardt W., Bohnert H.J.,
RA Bryant D.A.;
RT "Nucleotide sequence of the cyanelle DNA from Cyanophora paradoxa.";
RL Plant Mol. Biol. Rep. 13:327-332(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=LB555 / PRINGSHEIM;
RA Stirewalt V.L., Michalowski C.B., Loeffelhardt W., Bohnert H.J.,
RA Bryant D.A.;
RT "The complete sequence of the cyanelle genome of Cyanophora paradoxa:
RL the genetic complexity of a primitive plastid.";
RL (In) Schenk H.E.A., Herrmann R., Jeon K.W., Mueller N.E.,
RL Schwemmler W. (eds.);
RL Eukaryotism and Symbiosis, pp.40-48, Springer-Verlag, Heidelberg
RL (1997).
CC [1- SIMILARITY: BELONGS TO THE UPF0051 (YCF24) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U30821; AAA81250.1; -
DR Mendel; 7937; CYAPA; ycf24;1.
DR InterPro; IPR000825; UPF0051.
DR Pfam; PF01458; UPF0051; 1.
KW Cyanelle; Hypothetical protein.
SQ SEQUENCE 486 AA; 54267 MW; 3EC8520D61AB5C75 CRC64;

Query Match 21.4%; Score 6; DB 1; Length 486;
Best Local Similarity 100.0%; Pred. No. 34;
```

```

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
DB 312 TGSALT 317

RESULT 23
ID GLRK_CHICK STANDARD; PRT; 487 AA.
AC P19439;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Probable glutamate receptor precursor (kainate-binding protein).
GN KBR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Cerebellum;
RX MEDLINE=90081841; PubMed=2480525;
RA Gregor P., Maoz I., McKeown M., Teichberg V.I.;
RT "Molecular structure of the chick cerebellar kainate-binding subunit
of a putative glutamate receptor.";
RL Nature 342:689-692(1989).
CC -!- FUNCTION: L-GLUTAMATE ACTS AS AN EXCITATORY NEUROTRANSMITTER AT
MANY SYNAPSES IN THE CENTRAL NERVOUS SYSTEM. THE POSTSYNAPTIC
ACTIONS OF GLU ARE MEDIATED BY A VARIETY OF RECEPTORS THAT ARE
NAMED ACCORDING TO THEIR SELECTIVE AGONISTS.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL; X17700; CAA35693.1; -.
DR PIR; S07062; S07062.
DR HSP; P19491; 1GR2.
DR InterPro; IPR001320; Ion_glut_receptor.
DR InterPro; IPR001311; SBP_glut_receptor.
DR Pfam; PF00060; lig_chan; 1.
DR SMART; SM00079; PBPE; 1.
DR Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW Transmembrane.
FT SIGNAL 1 23
FT CHAIN 24 487 PROBABLE GLUTAMATE RECEPTOR.
FT DOMAIN 24 171 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 172 190 POTENTIAL.
FT TRANSMEM 207 225 POTENTIAL.
FT TRANSMEM 236 256 POTENTIAL.
FT TRANSMEM 420 440 POTENTIAL.
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 487 AA; 54355 MW; 45D71B2A94A27880 CRC64;

Query Match 21.4%; Score 6; DB 1; Length 487;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ENLPSL 16
DB 44 ENLPSL 49

RESULT 24
ID YC24_PORPU STANDARD; PRT; 487 AA.
AC P51240;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Hypothetical 54.4 kDa protein YCF24 (ORF487).
GN YCF24.
OS Porphyra purpurea.
OG Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
OX NCBI_TaxID=2787;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AVONPORT;
RA Reith M.E., Munnholland J.;
RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast
genome.";
RL Plant Mol. Biol. Rep. 13:333-335(1995).
CC -!- SIMILARITY: BELONGS TO THE UPF0051 (YCF24) FAMILY.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL; U38804; AAC08126.1; -.
DR Mendel; 10374; PORpu:ycf24; 1.
DR InterPro; IPR000825; UPF0051.
DR Pfam; PF01458; UPF0051; 1.
DR Chloroplast; Hypothetical protein.
SQ SEQUENCE 487 AA; 54396 MW; D892E16B1C84F384 CRC64;

Query Match 21.4%; Score 6; DB 1; Length 487;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
DB 313 TGSALT 318

RESULT 25
ID SUFB_ECOLI STANDARD; PRT; 495 AA.
AC P77522;
DT 15-JUL-1998 (Rel. 36, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sufb protein.
GN SUFB OR B1683.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA Sempel G., Seki Y., Sivasundaram S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
RN [3]
RP GENE NAME.
RC STRAIN=K12 / MG1655;
RX MEDLINE=99255563; PubMed=10322040;
RA Patzer S.I., Hantke K.;
RT "SufS is a Nifs-like protein, and SufD is necessary for stability of
RT the 2Fe-2S FnuF protein in Escherichia coli.";
RL J. Bacteriol. 181:3307-3309(1999).
CC -!- SIMILARITY: BELONGS TO THE UPF0051 (YCF24) FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AE000263; AAC74753.1; ALT_INIT.
DR EMBL; D90811; BAA15454.1; ALT_INIT.
DR EMBL; D90812; BAA15460.1; ALT_INIT.
DR EcoGene; EG13965; sufb.
DR InterPro; IPR000825; UPF0051.
DR Pfam; PF01458; UPF0051; 1.
KW Complete proteome.
SQ SEQUENCE 495 AA; 54745 MW; 3BB49EEADB6C1278 CRC64;
-----
Query Match 21.4%; Score 6; DB 1; Length 495;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGSALT 6
Db 321 TGSALT 326
|||||

RESULT 26
PVCC_PSEAE STANDARD; PRT; 500 AA.
AC O30372;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Poverdin chromophore biosynthetic protein pvcC.
GN PVCC OR PA2256.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=99315810; PubMed=10383985;
RA Stintzi A., Johnson Z., Stonehouse M., Ochsner U., Meyer J.M.,
RA Vasil M.L., Poole K.;
RT "The pvc gene cluster of Pseudomonas aeruginosa: role in synthesis of
RT the poverdine chromophore and regulation by PtxR and PvdS.";
RL J. Bacteriol. 181:4118-4124(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;

```

```

RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -!- COFACTOR: FAD (BY SIMILARITY).
CC -!- PATHWAY: PYOVERDINE CHROMOPHORE BIOSYNTHESIS.
CC -!- SIMILARITY: STRONG, TO P-HYDROXYPHENYLACETATE 3-HYDROXYLASES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AF002222; AAC21673.1; -.
DR EMBL; AE004651; AAG05644.1; -.
KW Oxidoreductase; FAD; Flavoprotein; Complete proteome.
FT CONFLICT 146 146 L -> F (IN REF. 1).
SQ SEQUENCE 500 AA; 55778 MW; 15612EB86C03EBBE CRC64;
-----
Query Match 21.4%; Score 6; DB 1; Length 500;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGSALT 6
Db 189 TGSALT 194
|||||

RESULT 27
MIS_RAT
ID MIS_RAT STANDARD; PRT; 553 AA.
AC P49000;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mullerian inhibiting factor precursor (MIS) (Anti-muellerian hormone)
DE (AMH) (Mullerian inhibiting substance).
DE AMH.
GN Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9241861; PubMed=1572639;
RA Haq C., Lee M.N., Tizard R., Wysk M., Demarinis J., Donahoe P.K.,
RA Cate R.L.;
RT "Isolation of the rat gene for Mullerian inhibiting substance.";
RL Genomics 12:665-669(1992).
CC -!- FUNCTION: THIS GLYCOPROTEIN, PRODUCED BY THE SERTOLI CELLS OF THE
CC TESTIS, CAUSES REGRESSION OF THE MUELLERIAN DUCT. IT ALSO IS ABLE
CC TO INHIBIT THE GROWTH OF TUMORS DERIVED FROM TISSUES OF MUELLERIAN
CC DUCT ORIGIN.
CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).

```

```

-----
CC EMBL; S98336; AAB22104.1; -.
DR InterPro: IPR002400; GF_cysknot.
DR InterPro: IPR001839; TGF-beta.
DR Pfam; PF00019; TGF-beta; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR ProDom; PD000357; TGF-beta; 1.
DR SMART; SM00204; TGF-beta; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Growth factor; Glycoprotein; Gonadal differentiation; Signal.
FT SIGNAL 1 21
FT PROPEP 22 ?
FT CHAIN ? 553
FT DISULFID 455 519
FT DISULFID 481 550
FT DISULFID 485 552
FT DISULFID 518 518
FT CARBOHYD 325 325
FT CARBOHYD 409 409
SQ SEQUENCE 553 AA; 58888 MW; 75DAF3949A038A69 CRC64;

Query Match 21.4%; Score 6; DB 1; Length 553;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
Db 428 QGLRAE 433
|||||

RESULT 28
ID MIS_MOUSE STANDARD; PRT; 555 AA.
AC P27106.
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Muellerian inhibiting factor precursor (MIS) (Anti-muellerian hormone)
DE (AMH) (Mullerian inhibiting substance).
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129; TISSUE=Testis;
RX MEDLINE=921146272; PubMed=1782869;
RA Muensterberg A., Lovell-Badge R.;
RT "Expression of the mouse anti-mullerian hormone gene suggests a role
in both male and female sexual differentiation.";
RL Development 113:613-624 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX MEDLINE=96081226; PubMed=8541848;
RA Dresser D.W., Hacker A., Lovell-Badge R., Guerrier D.;
RT "The genes for a spliceosomal protein (SAP62) and the anti-Mullerian
hormone (AMH) are contiguous.";
RL Hum. Mol. Genet. 4:1613-1618 (1995).
RN [3]
RP SEQUENCE OF 1-42 FROM N.A.
RC STRAIN=129;
RX MEDLINE=96081226; PubMed=8541848;
RA Dresser D.W., Hacker A., Lovell-Badge R., Guerrier D.;
RT "The genes for a spliceosomal protein (SAP62) and the anti-Mullerian
hormone (AMH) are contiguous.";
RL Hum. Mol. Genet. 4:1613-1618 (1995).
CC -1- FUNCTION: THIS GLYCOPROTEIN, PRODUCED BY THE SERTOLI CELLS OF THE
TESTIS, CAUSES REGRESSION OF THE MUELLERIAN DUCT. IT ALSO IS ABLE
TO INHIBIT THE GROWTH OF TUMORS DERIVED FROM TISSUES OF MUELLERIAN
DUCT ORIGIN.
CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.

```

```

-1- TISSUE SPECIFICITY: SERTOLI CELLS OF FETAL TESTES, AND TESTES
JUST AFTER BIRTH, BUT ABSENT IN ADULT TESTES. IN FEMALE, AMH
IS EXPRESSED AFTER BIRTH IN THE GRANULOSA CELLS OF THE FOLLICLE.
AMH EXPRESSION IS DEPENDENT ON THE DEGREE OF FOLLICULAR
MATURATION AND NOT ON THE AGE OF THE OVARY.
-1- MISCELLANEOUS: ALTHOUGH IT DOES NOT COMPETE WITH EGF FOR RECEPTOR
BINDING SITES, MIS CAN INHIBIT THE AUTOPHOSPHORYLATION OF THE EGF
RECEPTOR IN VITRO.
-1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; X63240; CAA44912.1; -.
DR EMBL; X83733; CAC10450.1; -.
DR PIR; S20100; S20100.
DR MGI; MGI:88006; Amh.
DR InterPro: IPR002400; GF_cysknot.
DR InterPro: IPR001839; TGF-beta.
DR Pfam; PF00019; TGF-beta; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR ProDom; PD000357; TGF-beta; 1.
DR SMART; SM00204; TGF-beta; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Growth factor; Glycoprotein; Gonadal differentiation; Signal.
FT SIGNAL 1 20
FT PROPEP 21 ?
FT CHAIN ? 555
FT DISULFID 457 521
FT DISULFID 483 552
FT DISULFID 487 554
FT DISULFID 520 520
FT CARBOHYD 62 62
FT CARBOHYD 326 326
FT CARBOHYD 410 410
FT SEQUENCE 555 AA; 59778 MW; D6A3A20C50306E29 CRC64;

Query Match 21.4%; Score 6; DB 1; Length 555;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
Db 429 QGLRAE 434
|||||

RESULT 29
ID MIS_BOVIN STANDARD; PRT; 575 AA.
AC P03972.
DT 23-OCT-1986 (Rel. 02, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Muellerian inhibiting factor precursor (MIS) (Anti-muellerian hormone)
DE (AMH) (Mullerian inhibiting substance).
GN AMH.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX MEDLINE=86218082; PubMed=3754790;
RA Cate R.I., Mattaliano R.J., Hession C., Tizard R., Farber N.M.,
RA Cheung A., Ninfa E.G., Frey A.Z., Gash D.J., Chow E.P., Fisher R.A.,
RA Bertonis J.M., Torres G., Wallner B.P., Ramachandran K.L.,

```

RA Ragin R.C., Manganaro T.F., McLaughlin D.T., Donahoe P.K.;
RT "Isolation of the bovine and human genes for Mullerian inhibiting
RL substance and expression of the human gene in animal cells.";
CC Cell 45:685-698(1986).
CC -1- FUNCTION: THIS GLYCOPROTEIN, PRODUCED BY THE SERPOLI CELLS OF THE
CC TESTIS, CAUSES REGRESSION OF THE MUELLERIAN DUCT. IT ALSO IS ABLE
CC TO INHIBIT THE GROWTH OF TUMORS DERIVED FROM TISSUES OF MUELLERIAN
CC DUCT ORIGIN.
CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC -1- MISCELLANEOUS: ALTHOUGH IT DOES NOT COMPETE WITH EGF FOR RECEPTOR
CC BINDING SITES, MIS CAN INHIBIT THE AUTOPHOSPHORYLATION OF THE EGF
CC RECEPTOR IN VITRO.
CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M13151; AAA98765.1; -;
DR PIR; A01398; WFBOM.
DR InterPro; IPR002400; GF_cysknot.
DR InterPro; IPR001839; TGF-beta.
DR Pfam; PF00019; TGF-beta; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR ProDom; PD000357; TGF-beta; 1.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Growth factor; Glycoprotein; Gonadal differentiation; Signal.
FT SIGNAL 1 17 POTENTIAL.
FT PROPEP 18 24 POTENTIAL.
FT CHAIN 25 575 MUELLERIAN INHIBITING FACTOR.
FT DISULFID 477 541 BY SIMILARITY.
FT DISULFID 503 572 BY SIMILARITY.
FT DISULFID 507 574 BY SIMILARITY.
FT DISULFID 540 540 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 78 78 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 344 344 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 575 AA; 60623 MW; 892B89C1AC8B5A8 CRC64;

Query Match 21.4%; Score 6; DB 1; Length 575;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
Db 447 QGLRAE 452
|||||

RESULT 30
HLI4_AERSA
ID HLY4_AERSA STANDARD; PRT; 578 AA.
AC Q08677;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hemolysin 4 precursor.
GN ASH4.
OS Aeromonas salmonicida.
OC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;
OC Aeromonas.
OX NCBI_TaxID=645;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=17-2;
RC MEDLINE=94142497; PubMed=8309354;
RX Hirono I., Aoki T.;
RA "Cloning and characterization of three hemolysin genes from Aeromonas
RT salmonicida.";

RL Microb. Pathog. 15:269-282(1993).
CC -1- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
CC CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
CC DEFINED.
CC -1- SIMILARITY: BELONGS TO THE AHH1/ASH4/HLYA/VVHA FAMILY OF
CC HEMOLYSINS.
CC -1- SIMILARITY: CONTAINS 1 RICIN B-TYPE LECTIN DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X65049; CAA46185.1; -;
DR InterPro; IPR001340; Hemlysn_pore.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00652; Ricin_B_lectin; 1.
DR SMART; SM00458; RICIN; 1.
DR PROSITE; PS00231; RICIN_B_LECTIN; 1.
KW Hemolysis; toxin; Lectin; Signal.
FT SIGNAL 1 ? POTENTIAL.
FT CHAIN 1 578 HEMOLYSIN 4.
FT DOMAIN 448 539 RICIN B-TYPE LECTIN.
SQ SEQUENCE 578 AA; 63400 MW; B07DB1FA854B8E52 CRC64;

Query Match 21.4%; Score 6; DB 1; Length 578;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
Db 367 QGLRAE 372
|||||

RESULT 31
XYNA_CALSR
ID XYNA_CALSR STANDARD; PRT; 684 AA.
AC P40944;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Endo-1,4-beta-xylanase A precursor (EC 3.2.1.8) (Xylanase A)
DE (1,4-beta-D-xylan xylanohydrolase A).
GN XYNA.
OS Caldicellulosiruptor sp. (strain Rt8B.4).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Thermoanaerobacter group; Caldicellulosiruptor.
OX NCBI_TaxID=28238;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=97077616; PubMed=8920183;
RA Dwivedi P.P., Gibbs M.D., Saul D.J., Bergquist P.L.;
RT "Cloning, sequencing and overexpression in Escherichia coli of a
RT xylanase gene, xyna from the thermophilic bacterium Rt8B.4 genus
RT Caldicellulosiruptor.";
RL Caldicellulosiruptor.
CC Appl. Microbiol. Biotechnol. 45:86-93(1996).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-xylosidic
CC linkages in xylans.
CC -1- PATHWAY: Xylan DEGRADATION.
CC -1- SIMILARITY: BELONGS TO CELLULASE FAMILY F (FAMILY 10 OF GLYCOSYL
CC HYDROLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

```

CC -----
DR EMBL; L18965; AAB42044.1; -.
DR PIR; S41788; S41788.
DR HSSP; P10478; 1XVZ.
DR InterPro; IPR003305; CBD_6.
DR InterPro; IPR001000; Glyco_hydro_10.
DR Pfam; PF00318; CBD_6; 2.
DR PRINTS; PR00134; GLYDRLASE10.
DR PROSITE; PS00591; GLYCOSYL_HYDROL_F10; 1.
KW Xylan degradation; Hydrolase; Glycosidase; Signal; Repeat.
FT SIGNAL 1 34 POTENTIAL.
FT CHAIN 35 684 ENDO-1,4-BETA-XYLANASE A.
FT ACT_SITE 490 490 PROTON DONOR (BY SIMILARITY).
FT ACT_SITE 598 598 NUCLEOPHILE (BY SIMILARITY).
SQ SEQUENCE 684 AA; 78353 MW; 0AE575F4PB4BA5E6 CRC64;

Query Match 21.4%; Score 6; DB 1; Length 684;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 NLPSLC 17
DB 352 NLPSLC 357

RESULT 32
CDGT_BACLI
ID CDGT_BACLI STANDARD; PRT; 718 AA.
AC P14014;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Cyclomaltodextrin glucanotransferase precursor (EC 2.4.1.19)
DE (Cyclodextrin-glycosyltransferase) (CGTase).
GN CGTA.
OS Bacillus licheniformis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1402;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90174931; PubMed=2137908;
RA Hill D.E., Aldape R., Rozzell J.D.;
RT "Nucleotide sequence of a cyclodextrin glucosyltransferase gene,
RT cgtA, from Bacillus licheniformis.";
RL Nucleic Acids Res. 18:199-199(1990).
CC -!- CATALYTIC ACTIVITY: Degrades starch to cyclodextrins by formation
CC of a 1,4-alpha-D-glucosidic bond.
CC -!- COFACTOR: BINDS TWO CALCIUM IONS.
CC -!- SUBUNIT: MONOMER.
CC -!- MISCELLANEOUS: CGTASE MAY CONSIST OF TWO PROTEIN DOMAINS: THE ONE
CC IN THE AMINO-TERMINAL SIDE CLEAVES THE ALPHA-1,4-GLUCOSIDIC BOND
CC IN STARCH, AND THE OTHER IN THE C-TERMINAL SIDE CATALYZES OTHER
CC ACTIVITIES, INCLUDING THE RECONSTITUTION OF AN
CC ALPHA-1,4-GLUCOSIDIC LINKAGE FOR CYCLIZING THE
CC MALTOOLIGOSACCHARIDE PRODUCED.
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15752; CAA33763.1; -.
DR PIR; S15920; ALBSMX.
DR HSSP; P30920; 1CGT.
DR InterPro; IPR000461; Alpha_amylase.

```

```

DR InterPro; IPR002044; CBD_4.
DR InterPro; IPR002909; IPT_TIG.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02806; alpha-amylase_C; 1.
DR Pfam; PF00886; CBD_4; 1.
DR Pfam; PF01833; TIG; 1.
DR ProDom; PD001568; CBD_4; 1.
KW Transferase; Glycosyltransferase; Calcium; Signal.
FT SIGNAL 1 34 BY SIMILARITY.
FT CHAIN 35 718 BY CYCLOMALTODEXTRIN GLUCANOTRANSFERASE.
FT DOMAIN 35 172 A1.
FT DOMAIN 173 236 B.
FT DOMAIN 237 440 A2.
FT DOMAIN 441 528 C.
FT DOMAIN 529 614 D.
FT DOMAIN 615 718 E.
FT DISULFID 77 84 BY SIMILARITY.
FT ACT_SITE 258 258 BY SIMILARITY.
FT ACT_SITE 291 291 BY SIMILARITY.
FT ACT_SITE 362 362 BY SIMILARITY.
SQ SEQUENCE 718 AA; 78002 MW; B3CDE14A81D5DC4E CRC64;

Query Match 21.4%; Score 6; DB 1; Length 718;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
DB 570 TGSALT 575

RESULT 33
BGAL_RHIME
ID BGAL_RHIME STANDARD; PRT; 755 AA.
AC Q59750;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Beta-galactosidase (EC 3.2.1.23) (Lactase).
GN LACZ.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=102F34;
RX MEDLINE=94215895; PubMed=8163182;
RA Fanning S., Leahy M., Sheehan D.;
RT "Nucleotide and deduced amino acid sequences of Rhizobium meliloti
RT 102F34 lacZ gene: comparison with prokaryotic beta-galactosidases and
RT human beta-glucuronidase.";
RL Gene 141:91-96(1994).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
CC galactose residues in beta-D-galactosides.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L20757; AAA26296.1; -.
DR InterPro; IPR001649; Glyco_hydro_2.
DR Pfam; PF00703; Glyco_hydro_2; 1.
DR Pfam; PF02836; Glyco_hydro_2_C; 1.
DR Pfam; PF02837; Glyco_hydro_2_N; 1.
DR PROSITE; PS00719; GLYCOSYL_HYDROL_F2_1; 1.
DR PROSITE; PS00608; GLYCOSYL_HYDROL_F2_2; FALSE_NEG.

```

KW Hydrolase; Glycosidase.
 FT ACT_SITE 382 382 PROTON DONOR (BY SIMILARITY).
 FT ACT_SITE 463 463 NUCLEOPHILE (BY SIMILARITY).
 SQ SEQUENCE 755 AA; 84148 MW; 09E339A92CC0D366 CRC64;

Query Match 21.4%; Score 6; DB 1; Length 755;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGSALT 6
 |||||
 Db 522 TGSALT 527

RESULT 34
 TFLG_HUMAN STANDARD; PRT; 1127 AA.
 AC Q9UPN9; O95855; O90J79;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Transcription intermediary factor 1-gamma (TIF1-gamma) (RET-fused gene
 DE 7 protein) (Rf97 protein).
 GN TIF1 OR RFG7 OR KIA1113.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99144725; PubMed=10022127;
 RA Venturini L., You J., Stadler M., Galien R., Lallemand V.,
 RA Koken M.H.M., Mattei M.-G., Ganser A., Chambon P., Losson R.,
 RA De The H.;
 RT "TIF1gamma, a novel member of the transcriptional intermediary factor
 RT 1 family.";
 RL Oncogene 18:1209-1217(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Brain;
 MEDLINE=99397452; PubMed=10470851;
 RA Kikuno R., Nagase T., Ishikawa K.-I., Hirose M., Miyajima N.,
 RA Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIV.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 6:197-205(1999).
 RN [3]
 RP SEQUENCE OF 76-1127 FROM N.A.
 RX TISSUE=Thyroid;
 MEDLINE=99367902; PubMed=10439047;
 RA Klugbauer S., Rabes H.;
 RT "The transcription coactivator htfl1 and a related protein are fused
 RT to the ret receptor tyrosine kinase in childhood papillary thyroid
 RT carcinomas.";
 RL Oncogene 18:4388-4393(1999).
 CC -!- FUNCTION: SEEMS TO ACT AS A TRANSCRIPTIONAL REPRESSOR.
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
 CC -!- SIMILARITY: CONTAINS 2 B BOX-TYPE ZINC FINGERS.
 CC -!- SIMILARITY: CONTAINS 1 BROMODOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 PHD-TYPE ZINC FINGER.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF119043; AAD17259.1; -

DR EMBL; AB029036; BAA83065.1; ALT_INIT.
 DR EMBL; AJ132948; CAB55313.1; -
 DR MW; 605769; -
 DR InterPro; IPR003649; Bbox_C.
 DR InterPro; IPR001487; Bromodomain.
 DR InterPro; IPR001965; PHD.
 DR InterPro; IPR000315; znf_box.
 DR InterPro; IPR001841; znf_ring.
 DR Pfam; PF00439; bromodomain; 1.
 DR Pfam; PF00628; PHD; 1.
 DR Pfam; PF00643; zf-B_box; 2.
 DR Pfam; PF00097; zf-C3HC4; 1.
 DR PRINTS; PR01406; BBOXZNFINGER.
 DR PRINTS; PR00503; BROMODOMAIN.
 DR SMART; SM00502; BBC; 1.
 DR SMART; SM00336; BBOX; 2.
 DR SMART; SM00297; BROMO; 1.
 DR SMART; SM00249; PHD; 2.
 DR SMART; SM00184; RING; 2.
 DR PROSITE; PS00633; BROMODOMAIN_1; FALSE_NEG.
 DR PROSITE; PS0014; BROMODOMAIN_2; 1.
 DR PROSITE; PS00119; ZF-BBOX; 2.
 DR PROSITE; PS00518; ZF-RING_1; 1.
 DR PROSITE; PS00089; ZF-RING_2; 1.
 KW Transcription regulation; Repressor; DNA-binding; Bromodomain;
 KW Zinc-finger; Nuclear protein; Repeat; Coiled coil.
 FT DOMAIN 13 18 POLY-GLY.
 FT DOMAIN 44 49 POLY-GLU.
 FT ZN_FING 125 154 RING-TYPE.
 FT ZN_FING 212 259 B BOX-TYPE 1.
 FT ZN_FING 271 312 B BOX-TYPE 2.
 FT DOMAIN 299 401 COILED COIL (POTENTIAL).
 FT DOMAIN 545 550 POLY-THR.
 FT ZN_FING 887 934 PHD-TYPE.
 FT DOMAIN 974 1046 BROMODOMAIN.
 FT CONFLICT 89 89 V -> E (IN REF. 3).
 FT CONFLICT 451 453 PAA -> LLH (IN REF. 3).
 FT CONFLICT 909 909 F -> S (IN REF. 3).
 FT CONFLICT 1037 1037 R -> T (IN REF. 1).
 FT CONFLICT 1114 1127 RKRLKSDRPVHIK -> OKTPKVR (IN REF. 1).
 SQ SEQUENCE 1127 AA; 122521 MW; 02E3880CFDCFDA3B CRC64;

Query Match 21.4%; Score 6; DB 1; Length 1127;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 ENLPSL 16
 |||||
 Db 675 ENLPSL 680

RESULT 35
 SIMA_DROME STANDARD; PRT; 1507 AA.
 AC Q24167; Q9VAA5;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Similar protein.
 GN SIMA OR CG7951.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96269413; PubMed=8682312;
 RA Nambu J.R., Chen W., Hu S., Crews S.T.;
 RT "The Drosophila melanogaster similar bHLH-PAS gene encodes a protein
 RT related to human hypoxia-inducible factor 1 alpha and Drosophila
 RT single-minded.";


```

RL Gene 172:249-254(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Brill J.F., Aghayani A., An H.-J., Andrews-Frannkoch C., Baldwin D.,
RA Ballev R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Beeson P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -!- FUNCTION: POSSIBLE DNA-BINDING TRANSCRIPTIONAL ACTIVATOR.
CC -!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
CC BHLH PROTEIN.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- TISSUE SPECIFICITY: UBQUITOUSLY EXPRESSED IN THE EMBRYO.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS. HIGHEST TO HIF-1 ALPHA.
CC -!- SIMILARITY: CONTAINS 2 PAS (PER-ARNT-SIM) DIMERIZATION DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U43090; AAC47303.1; -
DR EMBL; AF003772; AAF57008.2; -
DR FlyBase; FBgn0015542; sima.
DR InterPro; IPR003015; HLH_Myc.
DR InterPro; IPR001092; HLH_dlm.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAC.
DR Pfam; PF00785; PAC; 1.
DR Pfam; PF00989; PAS; 2.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.

```

```

DR PROSITE; PS00038; HELIX_LOOP_HELIX; 1.
DR PROSITE; PS0112; PAS; 2.
KW Repeat; DNA-binding; Nuclear protein; Transcription regulation;
KW Activator; Coiled coil.
FT DNA_BIND 72 85 BASIC DOMAIN.
FT DOMAIN 86 126 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT PAS 1. PAS 1.
FT DOMAIN 167 240 PAS 2.
FT DOMAIN 307 377 PAS 2.
FT DOMAIN 381 422 PAC.
FT DOMAIN 577 587 PRO-RICH.
FT DOMAIN 880 908 COILED COIL (POTENTIAL).
FT DOMAIN 982 1054 COILED COIL (POTENTIAL).
FT DOMAIN 1110 1162 COILED COIL (POTENTIAL).
FT DOMAIN 26 39 POLY-SER-.
FT DOMAIN 718 725 POLY-SER-.
FT DOMAIN 759 763 POLY-GLN.
FT DOMAIN 767 776 POLY-GLN.
FT DOMAIN 907 918 POLY-GLN.
FT DOMAIN 945 948 POLY-GLN.
FT DOMAIN 990 998 POLY-GLN.
FT DOMAIN 1020 1038 POLY-GLN.
FT DOMAIN 1113 1126 POLY-GLN.
FT DOMAIN 1146 1162 POLY-GLN.
FT DOMAIN 1205 1208 POLY-GLN.
FT DOMAIN 1277 1284 POLY-GLN.
FT DOMAIN 1298 1301 POLY-ASP.
FT CONFLICT 38 38 S -> A (IN REF. 1).
FT CONFLICT 345 345 S -> L (IN REF. 1).
FT CONFLICT 492 492 A -> V (IN REF. 1).
FT CONFLICT 588 588 T -> I (IN REF. 1).
FT CONFLICT 709 709 T -> K (IN REF. 1).
FT CONFLICT 776 776 Q -> QQQQ (IN REF. 1).
FT CONFLICT 895 895 Q -> QQ (IN REF. 1).
FT CONFLICT 902 902 G -> S (IN REF. 1).
FT CONFLICT 982 982 A -> T (IN REF. 1).
FT CONFLICT 1125 1126 MISSING (IN REF. 1).
FT CONFLICT 1154 1157 MISSING (IN REF. 1).
FT CONFLICT 1444 1444 F -> L (IN REF. 1).
FT CONFLICT 1447 1447 F -> C (IN REF. 1).
FT CONFLICT 1451 1451 S -> N (IN REF. 1).
FT CONFLICT 1494 1494 D -> G (IN REF. 1).
SQ SEQUENCE 1507 AA; 165824 MW; 4102939C8FBFB0C6 CRC64;

Query Match 21.4%; Score 6; DB 1; Length 1507;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LPSLCS 18
DB 825 LPSLCS 830
|||||

RESULT 36
HD_FUGRU STANDARD; PRT; 3148 AA.
AC P51112;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Huntingtin (Huntington's disease protein homolog) (HD protein).
GN HD.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95375788; PubMed=7647794;
RA Baxendale S., Abdulla S., Elgar G., Buck D., Berks M., Micklem G.,
RA Durbin R., Bates G., Brenner S., Beck S., Lehrach H.;

```

```

RT RT Huntington's disease genes." ;
RL Nat. Genet. 10:67-76(1995).
CC
CC -!- FUNCTION: MAY PLAY A ROLE IN MICROTUBULE-MEDIATED TRANSPORT OR
CC VESICLE FUNCTION.
CC
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- POLYMORPHISM: THE POLY-GLN REGION (FOUR RESIDUES) DOES NOT APPEAR
CC TO BE POLYMORPHIC, EXPLAINING THE ABSENCE OF A HD-LIKE DISORDER.
CC
CC -!- SIMILARITY: CONTAINS 10 HEAT REPEATS.
CC
CC -!- SIMILARITY: BELONGS TO THE HUNTINGTIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: X82939; CAA58112.1 ;
DR InterPro: IPR000357; HEAT_repeat.
DR InterPro: IPR000091; HUNTINGTIN.
DR PRINTS: PR00375; HUNTINGTIN.
DR PROSITE: PS50077; HEAT_REPEAT; 1.
KW Repeat.
FT DOMAIN 148 272 HEAT REPEATS DOMAIN 1.
FT DOMAIN 701 898 HEAT REPEATS DOMAIN 2.
FT DOMAIN 1527 1568 HEAT REPEATS DOMAIN 3.
FT DOMAIN 18 21 POLY-GLN.
FT DOMAIN 679 682 POLY-ALA.
FT DOMAIN 1104 1108 POLY-SER.
SQ SEQUENCE 3148 AA; 348932 MW; D9358676B0345243 CRC64;

Query Match 21.4%; Score 6; DB 1; Length 3148;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GURAE 26
Db 699 GURAE 704

RESULT 37
PSAL_BRAOL
ID PSAL_BRAOL STANDARD; PRT; 36 AA.
AC Q31909;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Photosystem I reaction center subunit VIII (PSI-I).
GN PSAL.
OS Brassica oleracea (Cauliflower).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3712;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Stange J.L., Dzelzkalns V.A.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY HELP IN THE ORGANIZATION OF THE PSAL SUBUNIT.
CC -!- SIMILARITY: BELONGS TO THE PSAL FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL: U34205; AAB05670.1; -.
DR InterPro: IPR001302; PSI_8.
DR Pfam: PF00796; PSI_8; 1.
DR ProDom: PD003995; PSI_8; 1.
KW Chloroplast; Photosystem I; Photosynthesis; Transmembrane.
FT TRANSMEM 8 28 POTENTIAL.
SQ SEQUENCE 36 AA; 3937 MW; BE9AECAL436086DF CRC64;

Query Match 17.9%; Score 5; DB 1; Length 36;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 NLPSSL 16
Db 5 NLPSSL 9

RESULT 38
PSAL_PEA
ID PSAL_PEA STANDARD; PRT; 40 AA.
AC P17227;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Photosystem I reaction center subunit VIII (PSI-I).
GN PSAL.
OS Pisum sativum (Garden pea).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. ALASKA;
RX MEDLINE=92224289; PubMed=1807835;
RA Nagano Y., Matsuno R., Sasaki Y.;
RT "Sequence and transcriptional analysis of the gene cluster trnQ-zfpa-
RL psal-ORF231-peta in pea chloroplasts." ;
RL Curr. Genet. 20:431-436(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. ALASKA;
RX MEDLINE=92005794; PubMed=1913879;
RA Smith A.G., Wilson R.J., Kaethner T.M., Willey D.L., Gray J.C.;
RT "Pea chloroplast genes encoding a 4 kDa polypeptide of photosystem I
RL and a putative enzyme of CI metabolism." ;
RL Curr. Genet. 19:403-410(1991).
RN [3]
RP SEQUENCE OF 1-16.
RX MEDLINE=90242987; PubMed=2185953;
RA Ikeuchi M., Hirano A., Hiyama T., Inoue Y.;
RT "Polypeptide composition of higher plant photosystem I complex.
RT Identification of psal, psaj and psak gene products." ;
RL FEBS Lett. 263:274-278(1990).
CC -!- FUNCTION: MAY HELP IN THE ORGANIZATION OF THE PSAL SUBUNIT.
CC -!- SIMILARITY: BELONGS TO THE PSAL FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: X56315; CAA39757.1; -.
DR EMBL: X54750; CAA38547.1; -.
DR PIR: S09730; S09730.
DR PIR: S14558; S14558.
DR PIR: S17921; S17921.

```

DR Mendel; 4428; PISsa;psai;1.
 DR InterPro; IPR001302; PSI_8.
 DR Pfam; PF00796; PSI_8; 1.
 DR ProDom; PD003995; PSI_8; 1.
 KW Chloroplast; Photosystem I; Photosynthesis; Transmembrane.
 FT TRANSMEM 8 28 POTENTIAL.
 SQ SEQUENCE 40 AA; 4469 MW; 3CE0B06A11A58E48 CRC64;

Query Match 17.9%; Score 5; DB 1; Length 40;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 NLP5L 16
 |||||
 Db 3 NLP5L 7

RESULT 39

GRK_BACCE STANDARD; PRT; 69 AA.
 AC Q9XBL1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycerate kinase (EC 2.7.1.31) (Fragment).
 GN GLXK.
 OS Bacillus cereus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 CC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1396;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 10987;
 RX MEDLINE=99231848; PubMed=10217496;
 RA Oksad O.A., Hegna I., Lindbaeck T., Rishovd A.L., Kolsto A.B.;
 RT "Genome organization is not conserved between Bacillus cereus and
 Bacillus subtilis.";
 RL Microbiology 145:621-631(1999).
 CC -!- CATALYTIC ACTIVITY: ATP + (R)-glycerate = ADP + 3-phospho-(R)-
 glycinate
 CC -!- SIMILARITY: BELONGS TO THE GLYCERATE KINASE FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

DR EMBL; AJ010129; CAB40583.1; -
 KW Transferase; Kinase.
 FT NON_TER 1
 SQ SEQUENCE 69 AA; 7502 MW; 7141E12A469878C5 CRC64;

Query Match 17.9%; Score 5; DB 1; Length 69;
 Best Local Similarity 100.0%; Pred. No. 68;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 AEENI 28
 |||||
 Db 46 AEENI 50

RESULT 40

APC4_RAT STANDARD; PRT; 88 AA.
 ID APC4_RAT
 AC P55797;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE Apolipoprotein C-IV precursor (Apo-CIV) (Apolipoprotein E-linked
 (ECL) (Fragment).
 GN APOC4 OR ECL.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92359549; PubMed=1379790;
 RA Shen P., Howlett G.J.;
 RT "Two coding regions closely linked to the rat apolipoprotein E gene:
 nucleotide sequences of rat apolipoprotein C-I and ECL cDNA.";
 RL Arch. Biochem. Biophys. 297:345-353(1992).
 CC -!- FUNCTION: MAY PARTICIPATE IN LIPOPROTEIN METABOLISM.
 CC -!- SUBCELLULAR LOCATION: Extracellular.
 CC -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A STOP CODON WAS
 READ THROUGH AT POSITION 70.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

DR EMBL; X56190; -; NOT_ANNOTATED.CDS
 KW Plasma; Lipid transport; Glycoprotein; Signal.
 FT SIGNAL 1 27 POTENTIAL.
 FT CHAIN 28 >88 APOLIPOPROTEIN C-IV.
 FT NON_TER 88
 SQ SEQUENCE 88 AA; 10245 MW; 6C02964F28E207D8 CRC64;

Query Match 17.9%; Score 5; DB 1; Length 88;
 Best Local Similarity 100.0%; Pred. No. 85;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LPSLC 17
 |||||
 Db 11 LPSLC 15

RESULT 41

HXA2_NOTVI STANDARD; PRT; 90 AA.
 ID HXA2_NOTVI
 AC F31261;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-JUL-1993 (Rel. 26, Last annotation update)
 DE Homeobox protein Hox-A2 (Nvbox-2.8) (Fragment).
 OS Notoththalmus viridescens (Eastern newt) (Triturus viridescens).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae;
 CC Notoththalmus.
 OX NCBI_TaxID=8316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92290273; PubMed=1351019;
 RA Belleville S., Beauchemin M., Tremblay M., Noiseux N., Savard P.;
 RT "Homeobox-containing genes in the newt are organized in clusters
 similar to other vertebrates.";
 RL Gene 114:179-186(1992).

CC -!- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
 A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH
 SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: BELONGS TO THE ANTP FAMILY OF HOMEBOX PROTEINS.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M84002; AAA49396.1; -
 DR PIR; JC1163; JC1163.
 DR HSSP; P14653; 1B72.
 DR InterPro; IPR001827; Antennapedia.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS00032; ANTENAPEDIA; PARTIAL.
 DR PROSITE; PS00071; HOMEBOX_2; 1.
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 KW Transcription regulation.
 FT NON_TER 1 1
 FT DNA_BIND 2 61 HOMEBOX.
 FT NON_TER 90 90
 SQ SEQUENCE 90 AA; 10764 MW; C9F486132E2B1FA7 CRC64;

Query Match 17.9%; Score 5; DB 1; Length 90;
 Best Local Similarity 100.0%; Pred. No. 87;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 8 QCKEN 12
 |||||

Db 62 QCKEN 66

RESULT 42

YDDX_ECOLI
 ID YDDX_ECOLI STANDARD; PRT; 92 AA.
 AC P76127; 1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein yddx.
 GN YDDX OR B1481.
 OS Escherichia coli.
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Escherichia.
 CC NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN-K12 / MG1655;
 RX MEDLINE-97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.
 RT "The complete genome sequence of Escherichia coli K-12."
 RL Science 277:1453-1474(1997).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE000245; AAC74554.1; -
 DR Ecogen; EG14301; yddx
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 92 AA; 10423 MW; 644091F41BF10E40 CRC64;

Query Match 17.9%; Score 5; DB 1; Length 92;
 Best Local Similarity 100.0%; Pred. No. 88;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 20 QGLRA 24
 |||||

Db 43 QGLRA 47

RESULT 43

A9_BRANA
 ID A9_BRANA STANDARD; PRT; 96 AA.
 AC O05772;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Tapetum-specific protein A9 precursor.
 GN A9.
 OS Brassica napus (Rape).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids II; Brassicales; Brassicaceae; Brassica.
 CC NCBI_TaxID=3708;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Anther;
 RX MEDLINE-92329721; PubMed=1627774;
 RA Paul W., Hodge R., Smartt S., Draper J., Scott R.;
 RT "The isolation and characterisation of the tapetum-specific
 Arabidopsis thaliana A9 gene."
 RL Plant Mol. Biol. 19:611-622(1992).
 CC -!- TISSUE SPECIFICITY: TAPETUM OF ANTHERS.
 CC -!- DEVELOPMENTAL STAGE: FOUND WHEN SPOROGENOUS CELLS ARE IN EARLY
 CC MEIOSIS. DISAPPEARS TOTALLY AS THE MICROSPORES GO INTO INTERPHASE,
 CC WHEN THE TAPETAL CELL LAYER DEGENERATES.
 CC -!- SIMILARITY: BELONGS TO THE A9 / FILI FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X61751; CAA43890.1; -
 DR EMBL; X61752; CAA43891.1; -
 DR PIR; S22467; S22467.
 DR InterPro; IPR003612; AAI.
 DR InterPro; IPR001768; Cereal_1ryp_amyl_inh.
 DR Pfam; PF00234; trypan-alpha-amyl; 1.
 DR SMART; SM00499; AAI; 1.
 KW Signal.
 FT SIGNAL 1 29 POTENTIAL.
 FT CHAIN 30 96 TAPETUM-SPECIFIC PROTEIN A9.
 FT DISULFID 32 70 BY SIMILARITY.
 FT DISULFID 42 59 BY SIMILARITY.
 FT DISULFID 60 85 BY SIMILARITY.
 FT DISULFID 72 92 BY SIMILARITY.
 SQ SEQUENCE 96 AA; 10276 MW; 5F14A3C13FBD9AD CRC64;

Query Match 17.9%; Score 5; DB 1; Length 96;
 Best Local Similarity 100.0%; Pred. No. 92;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 12 NLPSEL 16
 |||||

Db 86 NLPSEL 90

RESULT 44

Y508_MYCTU
 ID Y508_MYCTU STANDARD; PRT; 97 AA.
 AC Q11172;
 DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 10.7 kDa protein Rv0508.
 GN RV0508 OR MT0529 OR MTCY20G9.35.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=11773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekai F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z77162; CAB00934.1; -;
 DR EMBL; AE006953; AAK44752.1; -;
 DR TIGR; MT0529; -;
 DR Tuberculist; Rv0508; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 97 AA; 10710 MW; 8D9E498089799610 CRC64;

 Query Match 17.9%; Score 5; DB 1; Length 97;
 Best Local Similarity 100.0%; Pred. No. 93;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 21 GLRAE 25
 Db 50 GLRAE 54

 RESULT 45
 LV2G_HUMAN
 ID LV2G_HUMAN STANDARD; PRT; 111 AA.
 AC P01710;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig lambda chain V-II region BO.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE.
 RX MEDLINE=71103825; PubMed=5532228;
 RA Wikler M., Putnam F.W.;
 RT "Amino acid sequence of human lambda chains. 3. Tryptic peptides,
 chymotryptic peptides, and sequence of protein Bo.";
 RL J. Biol. Chem. 245:4488-4507(1970).
 CC -I- MISCELLANEOUS: THIS IS A BENICE-JONES PROTEIN.
 DR PIR; A01976; L2HUBO.
 DR HSSP; P01709; 2MCG.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_V.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IGV; 1.
 KW Immunoglobulin V region; Benice-Jones protein.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT DISULFID 22 90 BY SIMILARITY.
 FT NON_TER 111 111
 SQ SEQUENCE 111 AA; 11785 MW; 92F5A1BF72421BAC CRC64;

 Query Match 17.9%; Score 5; DB 1; Length 111;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 21 GLRAE 25
 Db 79 GLRAE 83

 RESULT 46
 YEVL_YEAST
 ID YEVL_YEAST STANDARD; PRT; 114 AA.
 AC P40076;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Hypothetical 12.9 kDa protein in SCS2-GLO3 intergenic region.
 GN YER121W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
 RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
 RA Chung E., Duncan M., Guzman E., Hartzell G., Hunkle-Smith S.,
 RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
 RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
 RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
 RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U18916; AAC03219.1; -;
 DR SGD; S0000923; YER121W.
 KW Hypothetical protein.
 SQ SEQUENCE 114 AA; 12941 MW; 71AEDD8607BE774B CRC64;

 Query Match 17.9%; Score 5; DB 1; Length 114;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 12 NLP5L 16

```
Db 75 NLP5L 79
|||||
RESULT 47
RL7A_THEAC
ID RL7A_THEAC STANDARD; PRT; 121 AA.
AC Q9HJ56;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L7Ae.
GN RPL7AE OR TALL16.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmales; Thermoplasmaceae;
OC Thermoplasma.
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mexes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RA "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum.";
RL Nature 407:508-513(2000).
CC -!- SIMILARITY: BELONGS TO THE L7AE FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL445066; CAC12243.1; -
CC InterPro; IPR004038; Ribosomal_L7Ae_L30e_S12e.
CC InterPro; IPR004037; Ribosomal_L7Ae.
CC Pfam; PF01248; Ribosomal_L7Ae; 1.
CC PRINTS; PR00881; L7ARSGFAMILY.
CC PROSITE; PS01082; RIBOSOMAL_L7AE; FALSE_NEG.
CC Ribosomal protein; Complete proteome.
CC SEQUENCE 121 AA; 13435 MW; 4156F6228E0E8B51 CRC64;

Query Match 17.9%; Score 5; DB 1; Length 121;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 LPSLC 17
|||||
Db 65 LPSLC 69

RESULT 48
APC4_RABIT
ID APC4_RABIT STANDARD; PRT; 124 AA.
AC P55057;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Apolipoprotein C-IV precursor (Apo-CIV) (PAP).
GN APC4.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=NEW ZEALAND; TISSUE=Liver;
RX MEDLINE=96139519; PubMed=8576182;
RA Zhang L., Kotite L., Havel R.J.;

"Identification, characterization, cloning, and expression of
RT apolipoprotein C-IV, a novel sialoglycoprotein of rabbit plasma
RT lipoproteins.";
RL J. Biol. Chem. 271:1776-1783(1996).
CC -!- FUNCTION: MAY PARTICIPATE IN LIPOPROTEIN METABOLISM.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: BLOOD PLASMA, ASSOCIATED PRIMARILY WITH VLDL
CC AND HDL. EXPRESSED MAINLY IN THE LIVER.
CC -!- PTM: SIALYLATED.
CC -!- PTM: PRESENT IN UP TO FIVE SIALYLATED ISOFORMS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U39356; AAC48513.1; -
CC Plasma; Lipid transport; Glycoprotein; VLDL; HDL; Signal.
CC SIGNAL 1 27
CC CHAIN 28 124 APOLIPOPROTEIN C-IV.
CC SEQUENCE 124 AA; 14113 MW; D4E42C51F078290C CRC64;

Query Match 17.9%; Score 5; DB 1; Length 124;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 LPSLC 17
|||||
Db 11 LPSLC 15

RESULT 49
JANA_DROME
ID JANA_DROME STANDARD; PRT; 135 AA.
AC P20348; Q9VAB6;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sex-regulated protein Janus-A.
GN JANA OR CG7933.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CANTON-S;
RX MEDLINE=89343970; PubMed=2503707;
RA Yanicostas C., Vincent A., Lepesant J.-A.;
RT "Transcriptional and posttranscriptional regulation contributes to
RT the sex-regulated expression of two sequence-related genes at the
RT Janus locus of Drosophila melanogaster.";
RL Mol. Cell. Biol. 9:2526-2535(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
```

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reihert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Massarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
CC -1- TISSUE SPECIFICITY: SOMATIC AND GERM LINE CELLS.
CC -1- DEVELOPMENTAL STAGE: THE NON-SEX-SPECIFIC TRANSCRIPT IS PRESENT AT
CC ALL STAGES.
CC -1- MISCELLANEOUS: TRANSCRIPTION OF JANA GIVES RISE TO TWO
CC SEX-SPECIFIC AND TO ONE NON-SEX-SPECIFIC TRANSCRIPTS.
CC -1- SIMILARITY: WITH SEX-REGULATED PROTEIN JANUS-B.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M27033; AAC34203.1; -;
DR EMBL; M27033; AAC34202.1; ALT_INIT.
DR EMBL; AE003772; AAF56997.1; -;
DR PIR; A32317; A32317.
DR FlyBase; FBgn0001280; jana.
SQ SEQUENCE 135 AA; 15220 MW; 2720237CE77F3132 CRC64;

Query Match 17.9%; Score 5; DB 1; Length 135;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GLRAE 25
Db 78 GLRAE 82
|||||

RESULT 50
RL23_YEAST
ID RL23_YEAST STANDARD; PRT; 137 AA.
AC P04451;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 60S ribosomal protein L23 (L17).
GN (RPL23A OR RPL17A OR YBL087C OR YBL0713) AND (RPL23B OR RPL17B OR
OS YER117W).
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;

RN SEQUENCE FROM N.A. (RPL23A).
RX MEDLINE=85014125; PubMed=6091033;
RA Leer R.J., van Raamsdonk-Duin M.M.C., Hagendoorn M.J.M., Mager W.H.,
RA Planta R.J.;
RT "Structural comparison of yeast ribosomal protein genes.";
RL Nucleic Acids Res. 12:6685-6700(1984).
RN [2]
RC SEQUENCE FROM N.A. (RPL23A).
RC STRAIN=S288C;
RX MEDLINE=96076635; PubMed=7502586;
RA Obermaier B., Gassenhuber J., Piravandi E., Domdey H.;
RT "Sequence analysis of a 78.6 kb segment of the left end of
RT Saccharomyces cerevisiae chromosome II.";
RL Yeast 11:1103-1112(1995).
RN [3]
RC SEQUENCE FROM N.A. (RPL23B).
RC STRAIN=S288C / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RC SEQUENCE FROM N.A. (RPL23B).
RX MEDLINE=95397593; PubMed=7668045;
RA Berroteran R.W., Hampsey M.;
RT "Sequence, map position and genome organization of the RPL17B gene,
RT encoding ribosomal protein L17b in *Saccharomyces cerevisiae*.";
RL Yeast 11:761-766(1995).
RN [5]
RP ERRATUM.
RA Berroteran R.W., Hampsey M.;
RL Yeast 12:91-91(1996).
CC -1- MISCELLANEOUS: THERE ARE TWO GENES FOR L23 IN YEAST.
CC -1- SIMILARITY: BELONGS TO THE L14P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01694; CAA25841.1; -;
DR EMBL; X79489; CAA56018.1; -;
DR EMBL; Z35848; CAA84908.1; -;
DR EMBL; U18916; AAC03215.1; -;
DR EMBL; U15653; AAA61906.1; -;
DR PIR; A02792; RSBY17.
DR HSSP; P04450; 1WHI.
DR SGD; S0000183; RPL23A.
DR SGD; S0000919; RPL23B.
DR InterPro; IPR000218; Ribosomal_L14.
DR Pfam; PF00238; Ribosomal_L14; 1.
DR ProDom; PD001093; Ribosomal_L14; 1.
DR PROSITE; PS00049; RIBOSOMAL_L14; 1.
KW Ribosomal protein; Multigene family.
SQ SEQUENCE 137 AA; 14473 MW; DEB983B3CB1DFAB1 CRC64;

Query Match 17.9%; Score 5; DB 1; Length 137;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GSAIT 6
Db 111 GSAIT 115
|||||

Search completed: August 22, 2002, 11:19:51
Job time: 178 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 11:20:34 ; Search time 42.29 seconds
(without alignments)
114.539 Million cell updates/sec

Title: US-09-856-199-2

Perfect score: 28

Sequence: 1 TGSAITXQCKENLPSCSXQGLRAEENI 28

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1200 summaries

Database :

SPTREMBL_19.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_invertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	39.3	380	11 Q9Z2T7	Q9Z2T7 mus musculus
2	7	25.0	370	16 Q9K554	Q9K554 bacillus ha
3	7	25.0	569	12 Q9IMX7	Q9IMX7 lumpy skin
4	7	25.0	817	4 Q9Y5D0	Q9Y5D0 homo sapien
5	7	25.0	932	4 Q9Y5G6	Q9Y5G6 homo sapien
6	6	21.4	94	2 O87046	O87046 vibrio chol
7	6	21.4	103	2 Q53981	Q53981 shigella dy
8	6	21.4	104	2 Q9RNC6	Q9RNC6 alcaligenes
9	6	21.4	104	2 Q9RAF1	Q9RAF1 ralstonia s
10	6	21.4	105	2 Q47965	Q47965 halomonadac
11	6	21.4	106	2 P74954	P74954 varlovorax
12	6	21.4	109	2 Q9RAD3	Q9RAD3 rhodofera
13	6	21.4	109	2 Q9T16	Q9T16 comamonas a
14	6	21.4	133	16 Q985Y1	Q985Y1 rhizobium l
15	6	21.4	141	16 Q98EG4	Q98EG4 rhizobium l
16	6	21.4	142	16 Q9CKW0	Q9CKW0 pasteurella

Q96mk6	homo sapien	148	4	Q96MK6	17	6	21.4
Q92z94	rhizobium m	153	16	Q92Z94	18	6	21.4
Q9mcd9	bacterioph	164	9	Q9MCD9	19	6	21.4
Q9w0v0	drosophila	166	5	Q9W0V0	20	6	21.4
Q9kf76	bacillus ha	174	16	Q9KFH6	21	6	21.4
Q30273	archaeoglob	175	17	Q30273	22	6	21.4
Q92ns5	rhizobium m	184	16	Q92NS5	23	6	21.4
P74447	synecocyst	186	16	P74447	24	6	21.4
Q9hw10	pseudomonas	187	16	Q9HWL0	25	6	21.4
Q99rn4	staphylococ	191	16	Q99RN4	26	6	21.4
Q38495	bacterioph	196	9	Q38495	27	6	21.4
Q9pi81	campylobact	210	16	Q9PI81	28	6	21.4
Q973x2	sulfolobus	222	17	Q973X2	29	6	21.4
Q9C661	arabidopsis	248	10	Q9C661	30	6	21.4
Q97fs0	clostridium	249	16	Q97FS0	31	6	21.4
Q93zvl	arabidopsis	258	10	Q93ZV1	32	6	21.4
P96452	rhizobium m	272	16	P96452	33	6	21.4
Q9kx47	alcaligenes	287	2	Q9KX47	34	6	21.4
Q93112	burkholderi	287	2	Q93112	35	6	21.4
P96312	burkholderi	288	2	P96312	36	6	21.4
Q9fgel	arabidopsis	301	10	Q9FGE1	37	6	21.4
Q12477	saccharomyc	304	3	Q12477	38	6	21.4
O30246	archaeoglob	324	17	O30246	39	6	21.4
Q9nwe9	homo sapien	332	4	Q9NWE9	40	6	21.4
Q9rdz8	legionella	339	2	Q9RDZ8	41	6	21.4
Q966m2	caenorhabdi	346	5	Q966M2	42	6	21.4
Q9xvu4	caenorhabdi	352	5	Q9XVU4	43	6	21.4
Q90634	gallus gall	359	13	Q90634	44	6	21.4
Q980u5	rhizobium l	366	16	Q980U5	45	6	21.4
Q96q98	homo sapien	377	4	Q96Q98	46	6	21.4
Q9n2h6	bos taurus	386	6	Q9N2H6	47	6	21.4
Q49722	mycobacteri	394	16	Q49722	48	6	21.4
Q987z2	rhizobium l	411	16	Q987Z2	49	6	21.4
P78356	homo sapien	416	4	P78356	50	6	21.4
O88377	rattus norv	416	11	O88377	51	6	21.4
Q9C6v9	arabidopsis	420	10	Q9C6V9	52	6	21.4
O17981	caenorhabdi	427	5	O17981	53	6	21.4
Q911b6	pseudomonas	443	16	Q911B6	54	6	21.4
Q9hql4	schizosacch	462	3	Q9HQL4	55	6	21.4
Q9b268	homo sapien	464	4	Q9B268	56	6	21.4
Q9b267	homo sapien	464	4	Q9B267	57	6	21.4
Q9vfe9	drosophila	469	5	Q9VFE9	58	6	21.4
P74623	synecocyst	481	16	P74623	59	6	21.4
Q9w214	drosophila	482	5	Q9W214	60	6	21.4
Q9pda3	xylella fas	485	16	Q9PDA3	61	6	21.4
Q9a762	caulobacter	489	16	Q9A762	62	6	21.4
Q92pk9	rhizobium m	489	16	Q92PK9	63	6	21.4
Q91wg2	mus musculu	490	11	Q91WG2	64	6	21.4
Q91t33	lumpy skin	494	12	Q91T33	65	6	21.4
Q9exp5	erwinia chr	499	2	Q9EXP5	66	6	21.4
Q9ne72	leishmania	505	5	Q9NE72	67	6	21.4
Q98nr7	rhizobium l	505	16	Q98NR7	68	6	21.4
O43558	homo sapien	508	4	O43558	69	6	21.4
Q90zif5	xenopus lae	509	13	Q90ZF5	70	6	21.4
Q99kn3	mus musculu	511	11	Q99KN3	71	6	21.4
Q9liq7	arabidopsis	521	10	Q9LIG7	72	6	21.4
Q92tj6	rhizobium m	537	16	Q92TJ6	73	6	21.4
Q9h5n1	homo sapien	538	4	Q9H5N1	74	6	21.4
Q94iv2	oryza sativ	544	10	Q94IV2	75	6	21.4
Q9zs97	arabidopsis	557	10	Q9ZS97	76	6	21.4
Q9h6k5	homo sapien	616	4	Q9H6K5	77	6	21.4
Q9awa7	chlorella s	625	10	Q9AWA7	78	6	21.4
Q9bx0	macaca fasc	655	6	Q9BGX0	79	6	21.4
Q915d9	pseudomonas	663	16	Q915D9	80	6	21.4
O30421	caldocellum	690	2	O30421	81	6	21.4
Q49498	mycoplasma	702	2	Q49498	82	6	21.4
Q9h7i3	homo sapien	721	4	Q9H7I3	83	6	21.4
Q9h7v5	homo sapien	735	4	Q9H7V5	84	6	21.4
Q9ndy0	leishmania	738	5	Q9NDY0	85	6	21.4
Q96st4	homo sapien	756	4	Q96ST4	86	6	21.4
Q961r9	homo sapien	767	4	Q961R9	87	6	21.4
Q9h7i7	homo sapien	780	4	Q9H7I7	88	6	21.4
Q9p2i7	homo sapien	796	4	Q9P2I7	89	6	21.4

90	6	21.4	810	10	Q9FRP8	Q9frp8 oryza sativ	163	5	17.9	66	12	Q91JD4	Q91jd4 hepatitis c
91	6	21.4	816	10	Q9LSJ0	Q9lsj0 arabidopsis	164	5	17.9	66	12	Q91JD3	Q91jd3 hepatitis c
92	6	21.4	819	3	Q59746	Q59746 schizosacch	165	5	17.9	66	12	Q91JD1	Q91jd1 hepatitis c
93	6	21.4	873	5	Q9VAH7	Q9vah7 drosophila	166	5	17.9	66	12	Q91JD0	Q91jd0 hepatitis c
94	6	21.4	896	4	Q75113	Q75113 homo sapien	167	5	17.9	66	12	Q91JC9	Q91jc9 hepatitis c
95	6	21.4	1083	4	Q9Y666	Q9y666 homo sapien	168	5	17.9	66	12	Q91JC8	Q91jc8 hepatitis c
96	6	21.4	1083	11	Q9GWL3	Q9gwl3 mus musculus	169	5	17.9	66	12	Q91JC7	Q91jc7 hepatitis c
97	6	21.4	1110	4	Q9C017	Q9c017 homo sapien	170	5	17.9	66	12	Q91JC6	Q91jc6 hepatitis c
98	6	21.4	1148	2	Q9FBR5	Q9fbr5 streptomyce	171	5	17.9	66	12	Q91JC5	Q91jc5 hepatitis c
99	6	21.4	1151	10	Q9CAP6	Q9cap6 arabidopsis	172	5	17.9	66	12	Q91JC4	Q91jc4 hepatitis c
100	6	21.4	1198	10	Q9C858	Q9c858 arabidopsis	173	5	17.9	66	12	Q91JC3	Q91jc3 hepatitis c
101	6	21.4	1202	10	Q9C858	Q9c858 arabidopsis	174	5	17.9	66	12	Q91JC2	Q91jc2 hepatitis c
102	6	21.4	1241	5	Q97004	Q97004 leishmania	175	5	17.9	66	12	Q91JC1	Q91jc1 hepatitis c
103	6	21.4	1297	5	Q9V604	Q9v604 drosophila	176	5	17.9	66	12	Q91JC0	Q91jc0 hepatitis c
104	6	21.4	1380	4	Q9P247	Q9p247 homo sapien	177	5	17.9	66	12	Q91JB9	Q91jb9 hepatitis c
105	6	21.4	1562	16	Q988W7	Q988w7 rhizobium 1	178	5	17.9	66	12	Q91JB7	Q91jb7 hepatitis c
106	6	21.4	1585	4	Q9P2D3	Q9p2d3 homo sapien	179	5	17.9	66	12	Q91JB5	Q91jb5 hepatitis c
107	6	21.4	1948	17	Q28189	Q28189 archaeoglob	180	5	17.9	66	12	Q91JB4	Q91jb4 hepatitis c
108	6	21.4	3010	12	Q9J3H8	Q9j3h8 hepatitis c	181	5	17.9	66	12	Q91JB3	Q91jb3 hepatitis c
109	6	21.4	3868	5	Q9W090	Q9w090 drosophila	182	5	17.9	66	12	Q91JB2	Q91jb2 hepatitis c
110	5	17.9	23	8	Q9T238	Q9t2j8 pisum sativ	183	5	17.9	66	12	Q91JB1	Q91jb1 hepatitis c
111	5	17.9	40	6	Q95L77	Q95l77 oryctolaqus	184	5	17.9	66	12	Q91JD3	Q91jd3 hepatitis c
112	5	17.9	44	6	Q9N1B0	Q9nlb0 monodelphis	185	5	17.9	66	12	Q91DD2	Q91dd2 hepatitis c
113	5	17.9	60	10	Q9U4L2	Q9u4l2 colpoda sp.	186	5	17.9	66	12	Q91DD1	Q91dd1 hepatitis c
114	5	17.9	60	10	Q9SMC5	Q9smc5 lycopersico	187	5	17.9	66	12	Q91DC9	Q91dc9 hepatitis c
115	5	17.9	66	12	Q9LJG2	Q9ljg2 hepatitis c	188	5	17.9	66	12	Q91DC8	Q91dc8 hepatitis c
116	5	17.9	66	12	Q9LJG1	Q9ljg1 hepatitis c	189	5	17.9	66	12	Q91DC7	Q91dc7 hepatitis c
117	5	17.9	66	12	Q9LJG0	Q9ljg0 hepatitis c	190	5	17.9	66	12	Q91DC6	Q91dc6 hepatitis c
118	5	17.9	66	12	Q9LJG9	Q9ljg9 hepatitis c	191	5	17.9	66	12	Q91DC5	Q91dc5 hepatitis c
119	5	17.9	66	12	Q9LJG8	Q9ljg8 hepatitis c	192	5	17.9	66	12	Q91DC4	Q91dc4 hepatitis c
120	5	17.9	66	12	Q9LJG7	Q9ljg7 hepatitis c	193	5	17.9	66	12	Q91DC3	Q91dc3 hepatitis c
121	5	17.9	66	12	Q9LJG6	Q9ljg6 hepatitis c	194	5	17.9	66	12	Q91DC2	Q91dc2 hepatitis c
122	5	17.9	66	12	Q9LJG5	Q9ljg5 hepatitis c	195	5	17.9	66	12	Q91DC1	Q91dc1 hepatitis c
123	5	17.9	66	12	Q9LJG4	Q9ljg4 hepatitis c	196	5	17.9	66	12	Q91DC0	Q91dc0 hepatitis c
124	5	17.9	66	12	Q9LJG3	Q9ljg3 hepatitis c	197	5	17.9	66	12	Q91DB9	Q91db9 hepatitis c
125	5	17.9	66	12	Q9LJG2	Q9ljg2 hepatitis c	198	5	17.9	66	12	Q91DB8	Q91db8 hepatitis c
126	5	17.9	66	12	Q9LJG1	Q9ljg1 hepatitis c	199	5	17.9	66	12	Q91DB6	Q91db6 hepatitis c
127	5	17.9	66	12	Q9LJG0	Q9ljg0 hepatitis c	200	5	17.9	66	12	Q91DB5	Q91db5 hepatitis c
128	5	17.9	66	12	Q9LJG9	Q9ljg9 hepatitis c	201	5	17.9	72	2	Q99PF8	Q99pf8 cricetus
129	5	17.9	66	12	Q9LJG8	Q9ljg8 hepatitis c	202	5	17.9	72	2	Q9AFL6	Q9afl6 shigella fl
130	5	17.9	66	12	Q9LJG7	Q9ljg7 hepatitis c	203	5	17.9	73	16	Q99X21	Q99x21 staphylococ
131	5	17.9	66	12	Q9LJG6	Q9ljg6 hepatitis c	204	5	17.9	74	4	Q9BGC3	Q9bgr3 homo sapien
132	5	17.9	66	12	Q9LJG5	Q9ljg5 hepatitis c	205	5	17.9	76	8	Q9TBS6	Q9tbs6 toxostoma l
133	5	17.9	66	12	Q9LJG4	Q9ljg4 hepatitis c	206	5	17.9	76	8	Q9TBS2	Q9tbs2 toxostoma r
134	5	17.9	66	12	Q9LJG3	Q9ljg3 hepatitis c	207	5	17.9	79	12	Q11361	Q11361 molluscum c
135	5	17.9	66	12	Q9LJG2	Q9ljg2 hepatitis c	208	5	17.9	81	13	Q9XVU5	Q9xvu5 ictalurus p
136	5	17.9	66	12	Q9LJG1	Q9ljg1 hepatitis c	209	5	17.9	82	4	Q13282	Q13282 homo sapien
137	5	17.9	66	12	Q9LJG0	Q9ljg0 hepatitis c	210	5	17.9	82	6	Q95280	Q95280 sus scrofa
138	5	17.9	66	12	Q9LJG9	Q9ljg9 hepatitis c	211	5	17.9	82	17	Q97UD0	Q97ud0 sulfolobus
139	5	17.9	66	12	Q9LJG8	Q9ljg8 hepatitis c	212	5	17.9	85	10	Q23224	Q23224 arabidopsis
140	5	17.9	66	12	Q9LJG7	Q9ljg7 hepatitis c	213	5	17.9	86	17	Q976F1	Q976f1 sulfolobus
141	5	17.9	66	12	Q9LJG6	Q9ljg6 hepatitis c	214	5	17.9	88	2	Q9XC08	Q9xc08 vibrio fisc
142	5	17.9	66	12	Q9LJG5	Q9ljg5 hepatitis c	215	5	17.9	88	9	Q80150	Q80150 bacterioph
143	5	17.9	66	12	Q9LJG4	Q9ljg4 hepatitis c	216	5	17.9	89	2	Q9ZHO4	Q9zho4 yersinia pe
144	5	17.9	66	12	Q9LJG3	Q9ljg3 hepatitis c	217	5	17.9	89	12	Q99226	Q99226 non-a non-b
145	5	17.9	66	12	Q9LJG2	Q9ljg2 hepatitis c	218	5	17.9	89	16	Q9PAT7	Q9pat7 xylella fas
146	5	17.9	66	12	Q9LJG1	Q9ljg1 hepatitis c	219	5	17.9	89	17	Q9HQJ5	Q9hqj5 thermoplasm
147	5	17.9	66	12	Q9LJG0	Q9ljg0 hepatitis c	220	5	17.9	89	17	Q979S5	Q979s5 thermoplasm
148	5	17.9	66	12	Q9LJG9	Q9ljg9 hepatitis c	221	5	17.9	91	4	Q13279	Q13279 homo sapien
149	5	17.9	66	12	Q9LJG8	Q9ljg8 hepatitis c	222	5	17.9	93	16	Q92JG0	Q92jg0 rickettsia
150	5	17.9	66	12	Q9LJG7	Q9ljg7 hepatitis c	223	5	17.9	94	2	P94197	P94197 agrobacteri
151	5	17.9	66	12	Q9LJG6	Q9ljg6 hepatitis c	224	5	17.9	95	11	Q99PU3	Q99pu3 rattus norv
152	5	17.9	66	12	Q9LJG5	Q9ljg5 hepatitis c	225	5	17.9	95	17	Q97BR3	Q97br3 thermoplasm
153	5	17.9	66	12	Q9LJG4	Q9ljg4 hepatitis c	226	5	17.9	96	5	O61725	O61725 eucidaris t
154	5	17.9	66	12	Q9LJG3	Q9ljg3 hepatitis c	227	5	17.9	96	8	Q9G8Y8	Q9g8y8 toxostoma r
155	5	17.9	66	12	Q9LJG2	Q9ljg2 hepatitis c	228	5	17.9	96	16	Q92FX5	Q92fx5 rickettsia
156	5	17.9	66	12	Q9LJG1	Q9ljg1 hepatitis c	229	5	17.9	99	10	Q9FIT0	Q9fit0 arabidopsis
157	5	17.9	66	12	Q9LJG0	Q9ljg0 hepatitis c	230	5	17.9	99	17	Q9HI23	Q9hiz3 thermoplasm
158	5	17.9	66	12	Q9LJG9	Q9ljg9 hepatitis c	231	5	17.9	101	2	Q9AQ32	Q9aq32 bradyrhizob
159	5	17.9	66	12	Q9LJG8	Q9ljg8 hepatitis c	232	5	17.9	101	4	Q9H534	Q9h534 homo sapien
160	5	17.9	66	12	Q9LJG7	Q9ljg7 hepatitis c	233	5	17.9	101	16	Q92I32	Q92i32 rickettsia
161	5	17.9	66	12	Q9LJG6	Q9ljg6 hepatitis c	234	5	17.9	102	5	Q9BH68	Q9bh68 drosophila
162	5	17.9	66	12	Q9LJG5	Q9ljg5 hepatitis c	235	5	17.9	103	12	Q68955	Q68955 hepatitis c

236	5	17.9	104	5	Q9BM25	Q9bmz5 drosophila	309	5	17.9	135	12	Q911C4	Q911c4 hepatitis c
237	5	17.9	104	5	Q9BM21	Q9bmz1 drosophila	310	5	17.9	135	12	Q911C3	Q911c3 hepatitis c
238	5	17.9	108	15	Q96985	Q96985 human immun	311	5	17.9	135	12	Q911C2	Q911c2 hepatitis c
239	5	17.9	108	15	Q96986	Q96986 human immun	312	5	17.9	135	12	Q911C1	Q911c1 hepatitis c
240	5	17.9	108	15	Q96987	Q96987 human immun	313	5	17.9	135	12	Q911C0	Q911b9 hepatitis c
241	5	17.9	108	15	Q96988	Q96988 human immun	314	5	17.9	135	12	Q911B9	Q911b9 hepatitis c
242	5	17.9	108	15	Q96989	Q96989 human immun	315	5	17.9	135	12	Q911B8	Q911b7 hepatitis c
243	5	17.9	109	5	Q9BM27	Q9bmz7 drosophila	316	5	17.9	135	12	Q911B7	Q911b7 hepatitis c
244	5	17.9	110	16	Q9PCP8	Q9pcr8 xylella fas	317	5	17.9	135	12	Q911B6	Q911b6 hepatitis c
245	5	17.9	111	4	Q9BRE8	Q9bre8 homo sapien	318	5	17.9	135	12	Q911B5	Q911b5 hepatitis c
246	5	17.9	111	10	Q94LR0	Q94lr0 oryza sativ	319	5	17.9	135	12	Q911B4	Q911b4 hepatitis c
247	5	17.9	111	17	Q9UZS8	Q9uzs8 pyrococcus	320	5	17.9	135	12	Q911B3	Q911b3 hepatitis c
248	5	17.9	112	16	Q989N2	Q989n2 rhizobium 1	321	5	17.9	135	12	Q911B2	Q911b2 hepatitis c
249	5	17.9	113	13	Q42461	Q42461 squalus aka	322	5	17.9	135	12	Q911B1	Q911b1 hepatitis c
250	5	17.9	113	13	Q91BH9	Q91bh9 squalus aka	323	5	17.9	135	12	Q911B0	Q911b0 hepatitis c
251	5	17.9	113	15	Q9ELU3	Q9elu3 human immun	324	5	17.9	135	12	Q911A9	Q911a9 hepatitis c
252	5	17.9	115	4	Q9BZK2	Q9bzk2 homo sapien	325	5	17.9	135	12	Q911A8	Q911a8 hepatitis c
253	5	17.9	115	4	Q9UG71	Q9ug71 homo sapien	326	5	17.9	135	12	Q911A7	Q911a7 hepatitis c
254	5	17.9	116	5	Q9BNB7	Q9bnb7 calyptophth	327	5	17.9	135	12	Q911A6	Q911a6 hepatitis c
255	5	17.9	117	2	Q9FB32	Q9fb32 streptomyce	328	5	17.9	136	2	Q9RC70	Q9rc70 bacillus ha
256	5	17.9	117	16	Q99SE8	Q99se8 staphylococ	329	5	17.9	136	2	Q9KZ93	Q9kz93 streptomyce
257	5	17.9	117	17	Q59463	Q59463 pyrococcus	330	5	17.9	136	4	Q9HBR5	Q9hbr5 homo sapien
258	5	17.9	118	4	Q9NQ99	Q9nqn9 homo sapien	331	5	17.9	136	16	P74444	P74444 synechocyst
259	5	17.9	118	14	Q99IU5	Q99iu5 uncultured	332	5	17.9	137	4	Q95653	Q95653 homo sapien
260	5	17.9	118	17	Q9YCB8	Q9ycb8 aeropyrum p	333	5	17.9	137	5	Q9U6R0	Q9u6r0 toxoplasma
261	5	17.9	119	9	Q38076	Q38076 bacterioph	334	5	17.9	138	5	Q95T66	Q95t66 drosophila
262	5	17.9	120	16	Q9CF83	Q9cf83 lactococcus	335	5	17.9	139	4	Q96MD1	Q96md1 homo sapien
263	5	17.9	121	2	Q9RNU0	Q9rnu0 solar lake	336	5	17.9	140	5	Q9WLY7	Q9wly7 drosophila
264	5	17.9	121	2	Q9RMR3	Q9rmr3 solar lake	337	5	17.9	140	5	Q9GNE2	Q9gne2 aedes aegyp
265	5	17.9	121	2	Q9AFR4	Q9afr4 shigella fl	338	5	17.9	140	5	Q962Y9	Q962y9 spodoptera
266	5	17.9	121	17	Q97BK8	Q97bk8 thermoplasm	339	5	17.9	140	6	Q9X5U3	Q9x5u3 canis famil
267	5	17.9	122	11	Q9D8R7	Q9d8r7 mus musculu	340	5	17.9	140	10	Q9AV77	Q9av77 oryza sativ
268	5	17.9	124	4	Q9NZD9	Q9nzd9 homo sapien	341	5	17.9	140	10	Q9ATF6	Q9atf6 castanea sa
269	5	17.9	124	16	Q9KG60	Q9kg60 bacillus ha	342	5	17.9	140	11	Q9PCQ4	Q9pcq4 mus musculu
270	5	17.9	126	2	Q9RKC9	Q9rk9 streptomyce	343	5	17.9	140	11	Q9CZ66	Q9cze6 mus musculu
271	5	17.9	126	17	Q9YEE3	Q9yee3 aeropyrum p	344	5	17.9	140	13	Q9W6M3	Q9w6m3 dicentrarch
272	5	17.9	129	2	Q9KX79	Q9kx79 streptomyce	345	5	17.9	140	13	Q9DFL3	Q9dfl3 gillichthys
273	5	17.9	130	2	Q9R2P5	Q9r2p5 pectobacter	346	5	17.9	141	5	Q9V5K1	Q9v5k1 drosophila
274	5	17.9	130	2	Q9XC13	Q9xc13 shigella fl	347	5	17.9	141	11	Q9CUU2	Q9cuu2 mus musculu
275	5	17.9	130	2	Q9RB06	Q9rb06 pectobacter	348	5	17.9	141	11	Q99JV7	Q99jv7 mus musculu
276	5	17.9	130	11	Q9CVV1	Q9cvv1 mus musculu	349	5	17.9	142	4	Q9H3F3	Q9h3f3 homo sapien
277	5	17.9	133	2	Q9F9X1	Q9f9x1 burkholderi	350	5	17.9	142	11	P70621	P70621 rattus norv
278	5	17.9	133	9	Q21906	Q21906 bacterioph	351	5	17.9	143	13	Q9YHX0	Q9yhx0 dicentrarch
279	5	17.9	133	10	Q65068	Q65068 picea maria	352	5	17.9	144	2	Q32569	Q32569 escherichia
280	5	17.9	133	11	Q99386	Q99386 mus musculu	353	5	17.9	145	4	Q9UL20	Q9ul20 homo sapien
281	5	17.9	134	4	Q9BTQ7	Q9btq7 homo sapien	354	5	17.9	146	16	Q98EF5	Q98ef5 rhizobium 1
282	5	17.9	134	17	Q970Q6	Q970q6 sulfolobus	355	5	17.9	147	2	P71267	P71267 escherichia
283	5	17.9	135	4	Q99779	Q99779 homo sapien	356	5	17.9	147	2	P71268	P71268 escherichia
284	5	17.9	135	5	Q9BH48	Q9bh48 drosophila	357	5	17.9	147	2	P71270	P71270 escherichia
285	5	17.9	135	5	Q9BM98	Q9bm98 drosophila	358	5	17.9	147	2	P71272	P71272 escherichia
286	5	17.9	135	5	Q9BM97	Q9bm97 drosophila	359	5	17.9	147	16	Q97NS3	Q97ns3 streptococ
287	5	17.9	135	5	Q9BM96	Q9bm96 drosophila	360	5	17.9	149	10	Q9FM74	Q9fm74 arabidopsis
288	5	17.9	135	5	Q9BM99	Q9bm99 drosophila	361	5	17.9	150	3	Q9UU32	Q9uu32 schizosacch
289	5	17.9	135	5	Q95VY7	Q95vy7 drosophila	362	5	17.9	150	16	Q9HUE3	Q9hue3 pseudomonas
290	5	17.9	135	10	Q9SMI7	Q9smi7 cyanophora	363	5	17.9	152	2	Q9ADL7	Q9adl7 polyangium
291	5	17.9	135	12	Q911E3	Q911e3 hepatitis c	364	5	17.9	153	2	Q98AM0	Q98am0 rhizobium 1
292	5	17.9	135	12	Q911E2	Q911e2 hepatitis c	365	5	17.9	153	16	Q9R9P7	Q9r9p7 aeromonas p
293	5	17.9	135	12	Q911E1	Q911e1 hepatitis c	366	5	17.9	154	2	Q9R9P7	Q9r9p7 aeromonas p
294	5	17.9	135	12	Q911E0	Q911e0 hepatitis c	367	5	17.9	154	3	Q94132	Q94132 piromyces s
295	5	17.9	135	12	Q911D9	Q911d9 hepatitis c	368	5	17.9	154	3	Q90UP7	Q90up7 piromyces s
296	5	17.9	135	12	Q911D8	Q911d8 hepatitis c	369	5	17.9	154	16	Q92HC9	Q92hc9 rickettsia
297	5	17.9	135	12	Q911D7	Q911d7 hepatitis c	370	5	17.9	155	5	O02421	O02421 bdelloura c
298	5	17.9	135	12	Q911D6	Q911d6 hepatitis c	371	5	17.9	155	16	Q9CLZ6	Q9clz6 pasteurella
299	5	17.9	135	12	Q911D4	Q911d4 hepatitis c	372	5	17.9	156	10	Q9FJY2	Q9fjy2 arabidopsis
300	5	17.9	135	12	Q911D3	Q911d3 hepatitis c	373	5	17.9	157	10	O22686	O22686 arabidopsis
301	5	17.9	135	12	Q911D2	Q911d2 hepatitis c	374	5	17.9	158	16	Q9CCM4	Q9ccm4 mycobacteri
302	5	17.9	135	12	Q911D1	Q911d1 hepatitis c	375	5	17.9	159	16	Q98M04	Q98m04 rhizobium 1
303	5	17.9	135	12	Q911D0	Q911d0 hepatitis c	376	5	17.9	159	10	Q92VC7	Q92vc7 arabidopsis
304	5	17.9	135	12	Q911C9	Q911c9 hepatitis c	377	5	17.9	160	10	Q949E5	Q949e5 oryza sativ
305	5	17.9	135	12	Q911C8	Q911c8 hepatitis c	378	5	17.9	161	4	Q9HQ3	Q9hq3 homo sapien
306	5	17.9	135	12	Q911C7	Q911c7 hepatitis c	379	5	17.9	162	2	Q03642	Q03642 escherichia
307	5	17.9	135	12	Q911C6	Q911c6 hepatitis c	380	5	17.9	162	16	Q92LP3	Q92lp3 rhizobium m
308	5	17.9	135	12	Q911C5	Q911c5 hepatitis c	381	5	17.9	162	17	O30324	O30324 archaeoglob

674	5	17.9	310	5	Q9TW87	Q9tw87 drosophila	747	16	Q9HYT1	Q9hyt1 pseudomonas
675	5	17.9	310	5	Q9TVG2	Q9tv92 drosophila	748	5	O02586	O02586 spirometra
676	5	17.9	310	10	Q9FJP2	Q9fjp2 arabidopsis	749	5	Q95NF1	Q95nf1 spirometra
677	5	17.9	310	10	Q9FRV4	Q9frv4 arabidopsis	750	5	Q9YMH4	Q9ymh4 lymantria d
678	5	17.9	310	10	Q9M4B6	Q9m4b6 arabidopsis	751	5	P72614	P72614 synecocyst
679	5	17.9	310	16	Q97L57	Q97l57 clostridium	752	5	P72614	P72614 synecocyst
680	5	17.9	311	2	Q9JP44	Q9jp44 rhizobium m	753	5	Q9W4M1	Q9w4m1 drosophila
681	5	17.9	311	13	Q99162	Q99162 xiphophorus	754	5	Q9LIL8	Q9lil8 arabidopsis
682	5	17.9	311	15	Q98395	Q98395 mouse mamma	755	5	P73759	P73759 synecocyst
683	5	17.9	314	2	Q9S5A2	Q9s5a2 listeria mo	756	5	Q96KT7	Q96kt7 homo sapien
684	5	17.9	314	16	Q92BP0	Q92bp0 listeria in	757	5	Q9Z954	Q9z954 chlamydia p
685	5	17.9	314	16	Q92AD4	Q92ad4 listeria in	758	5	Q9L9N2	Q9l9n2 borrelia bu
686	5	17.9	315	2	Q9RL57	Q9rl57 streptomyce	759	5	Q9ZBX7	Q9zbx7 streptomyce
687	5	17.9	315	10	Q94DN2	Q94dn2 oryza sativ	760	5	Q9D866	Q9d866 homo sapien
688	5	17.9	315	11	Q62473	Q62473 mus musculu	761	5	Q9WV03	Q9wv03 mus musculu
689	5	17.9	315	11	Q64286	Q64286 mus musculu	762	5	Q9L784	Q9l784 clostridium
690	5	17.9	315	11	Q61911	Q61911 mus musculu	763	5	O05417	O05417 clostridium
691	5	17.9	315	11	Q64354	Q64354 mus musculu	764	5	Q9NXZ4	Q9nxz4 homo sapien
692	5	17.9	315	13	Q9W600	Q9w600 gallus gall	765	5	Q9CXD6	Q9cxd6 mus musculu
693	5	17.9	315	15	Q91ZT9	Q91zt9 endogenous	766	5	O16040	O16040 drosophila
694	5	17.9	315	15	Q83411	Q83411 mouse mamma	767	5	O16040	O16040 drosophila
695	5	17.9	315	16	Q9KAC6	Q9kac6 bacillus ha	768	5	O86377	O86377 neisseria a
696	5	17.9	316	16	Q9RY95	Q9ry95 deinoococcus	769	5	Q9DWD1	Q9dwd1 rat cytomeg
697	5	17.9	317	15	Q86679	Q86679 mouse mamma	770	5	Q9ZWR9	Q9zwr9 rhizobium m
698	5	17.9	317	15	Q98343	Q98343 mouse mamma	771	5	O00421	O00421 homo sapien
699	5	17.9	317	16	Q91264	Q91264 pseudomonas	772	5	O00421	O00421 homo sapien
700	5	17.9	317	16	Q9H2X2	Q9hzx2 pseudomonas	773	5	Q9UPG0	Q9upg0 homo sapien
701	5	17.9	318	10	Q93936	Q93936 betula verr	774	5	Q96KP5	Q96kp5 homo sapien
702	5	17.9	318	16	Q99ZWR	Q99zwr streptococc	775	5	O917S2	O917s2 drosophila
703	5	17.9	318	16	Q92QR6	Q92qr6 rhizobium m	776	5	O44901	O44901 caenorhabdi
704	5	17.9	319	5	O44369	O44369 bactocera	777	5	Q9XSD7	Q9xsd7 macaca mula
705	5	17.9	319	12	Q68699	Q68699 hepatitis c	778	5	Q9LLX4	Q9llx4 brassica na
706	5	17.9	319	12	Q68700	Q68700 hepatitis c	779	5	Q9L7985	Q9l7985 archaeoglob
707	5	17.9	319	12	Q68701	Q68701 hepatitis c	780	5	Q97RN7	Q97rn7 streptococc
708	5	17.9	319	15	Q9IZT4	Q9izt4 exogenous m	781	5	Q9Z669	Q9z669 zymomonas m
709	5	17.9	320	11	Q99L42	Q99l42 mus musculu	782	5	O9C4N5	O9c4n5 corynebact
710	5	17.9	320	11	Q61912	Q61912 mus musculu	783	5	Q9RVT6	Q9rvt6 deinoococcus
711	5	17.9	320	15	Q83409	Q83409 mouse mamma	784	5	Q9XIA5	Q9xia5 arabidopsis
712	5	17.9	320	15	Q86678	Q86678 mouse mamma	785	5	Q9LUC0	Q9luc0 arabidopsis
713	5	17.9	320	15	Q9WHW0	Q9whw0 mouse mamma	786	5	Q9SNK8	Q9snk8 oryza sativ
714	5	17.9	321	2	Q9S5F8	Q9s5f8 escherichia	787	5	Q9L1X0	Q9l1x0 streptomyce
715	5	17.9	321	16	Q9CEC5	Q9cec5 lactococcus	788	5	Q9YMU2	Q9ymu2 lymantria d
716	5	17.9	322	4	Q9I5Y7	Q9i5y7 homo sapien	789	5	Q9X8I6	Q9x8i6 streptomyce
717	5	17.9	322	15	Q9UNF4	Q9unf4 homo sapien	790	5	Q9VAF3	Q9vaf3 drosophila
718	5	17.9	322	15	Q83396	Q83396 mouse mamma	791	5	Q9CX55	Q9cx55 mus musculu
719	5	17.9	322	15	Q89534	Q89534 mouse mamma	792	5	O40366	O40366 medicago sa
720	5	17.9	322	16	P71065	P71065 bacillus su	793	5	Q9YMT2	Q9ymt2 lymantria d
721	5	17.9	322	16	O25949	O25949 helicobacte	794	5	O98E13	O98e13 rhizobium l
722	5	17.9	322	16	Q9ZJ87	Q9zj87 helicobacte	795	5	Q9QGP2	Q9qgp2 equine infe
723	5	17.9	323	2	O85340	O85340 escherichia	796	5	Q9QGP2	Q9qgp2 equine infe
724	5	17.9	324	2	Q9Z178	Q9zi78 lactococcus	797	5	O05654	O05654 bacillus ce
725	5	17.9	324	5	Q9NHE2	Q9nhe2 leishmania	798	5	P78828	P78828 schizosach
726	5	17.9	324	15	Q83390	Q83390 exogenous m	799	5	Q9LCY9	Q9lcy9 oryza sativ
727	5	17.9	324	16	P73405	P73405 synecocyst	800	5	O9HY04	O9hy04 pseudomonas
728	5	17.9	324	16	Q9CF05	Q9cf05 lactococcus	801	5	O28416	O28416 archaeoglob
729	5	17.9	325	15	Q83412	Q83412 mouse mamma	802	5	O75307	O75307 homo sapien
730	5	17.9	326	13	Q90771	Q90771 gallus gall	803	5	O18082	O18082 homo sapien
731	5	17.9	327	10	Q9SXA1	Q9sxv1 nicotiana t	804	5	Q96RZ0	Q96rz0 homo sapien
732	5	17.9	327	17	Q9YAD7	Q9yad7 aeropyrum p	805	5	O18082	O18082 caenorhabdi
733	5	17.9	328	16	Q98GD7	Q98gd7 rhizobium l	806	5	Q965X7	Q965x7 caenorhabdi
734	5	17.9	329	17	Q9V2M7	Q9v2m7 pyrococcus	807	5	O9LET9	O9let9 arabidopsis
735	5	17.9	330	10	Q9W8K2	Q9w8k2 arabidopsis	808	5	Q98N29	Q98nz9 rhizobium l
736	5	17.9	330	11	Q9CY60	Q9cy60 mus musculu	809	5	Q9RLB4	Q9rlb4 rickettsia
737	5	17.9	331	10	Q93V84	Q93v84 arabidopsis	810	5	Q9H9W7	Q9h9w7 homo sapien
738	5	17.9	332	2	O69946	O69946 streptomyce	811	5	O91BW8	O91bw8 helicoverpa
739	5	17.9	332	16	Q97RW9	Q97rw9 streptococc	812	5	O9DG23	O9dg23 gallus gall
740	5	17.9	333	2	O68566	O68566 xanthomonas	813	5	O9Y6I5	O9y6i5 homo sapien
741	5	17.9	333	3	Q9XAW7	Q9xaw7 pseudomonas	814	5	O9FXA0	O9fxa0 arabidopsis
742	5	17.9	333	3	Q9UWE8	Q9uwe8 yarrowia li	815	5	O94K02	O94k02 arabidopsis
743	5	17.9	334	4	Q96FK1	Q96fk1 homo sapien	816	5	O9XBG6	O9xbg6 bradyrhizob
744	5	17.9	334	6	Q95264	Q95264 sus scrofa	817	5	O984F1	O984f1 rhizobium l
745	5	17.9	334	16	Q98B15	Q98b15 rhizobium l	818	5	Q9S0E7	Q9s0e7 borrelia bu
746	5	17.9	334	16	Q92DC3	Q92dc3 listeria in	819	5	Q9DEZ0	Q9dez0 brachydanio

820	5	17.9	364	10	Q39052	Q39052 arabisdopsis	893	5	17.9	398	5	044370	044370 bactrocera
821	5	17.9	366	2	Q93S72	Q93S72 aminobacter	894	5	17.9	399	5	Q39HA3	Q39HA3 streptomyce
822	5	17.9	366	2	Q93K07	Q93K07 lactobacilli	895	5	17.9	399	5	Q17840	Q17840 caenorhabdi
823	5	17.9	366	16	Q9HYM3	Q9HYM3 pseudomonas	896	5	17.9	400	10	Q9SRM3	Q9SRM3 arabisdopsis
824	5	17.9	367	3	Q9PE17	Q9PE17 schizosacch	897	5	17.9	400	13	Q91703	Q91703 xenopus lae
825	5	17.9	368	2	Q9KZ85	Q9KZ85 streptomyce	898	5	17.9	400	13	Q73818	Q73818 xenopus lae
826	5	17.9	368	4	Q9H841	Q9H841 homo sapien	899	5	17.9	400	13	Q90YD6	Q90YD6 xenopus tro
827	5	17.9	368	5	Q9VEEA	Q9VEEA drosophila	900	5	17.9	400	16	Q9PGE6	Q9PGE6 xylella fas
828	5	17.9	368	14	Q91US9	Q91US9 plasmid psb	901	5	17.9	401	5	Q965Y8	Q965Y8 caenorhabdi
829	5	17.9	368	16	Q97SZ5	Q97SZ5 streptococc	902	5	17.9	401	16	Q9KLO2	Q9KLO2 vibrio chol
830	5	17.9	369	13	Q9DEZ2	Q9DEZ2 xenopus lae	903	5	17.9	402	3	Q9P870	Q9P870 agaricus bi
831	5	17.9	369	16	Q9RZ18	Q9RZ18 deinococcus	904	5	17.9	402	10	Q9LXN2	Q9LXN2 arabisdopsis
832	5	17.9	370	4	Q75059	Q75059 homo sapien	905	5	17.9	402	10	Q9FY63	Q9FY63 arabisdopsis
833	5	17.9	370	10	Q9LGV4	Q9LGV4 oryza sativ	906	5	17.9	402	12	Q9JFR8	Q9JFR8 wheat roset
834	5	17.9	370	16	Q92JRH	Q92JRH rhizobium m	907	5	17.9	402	16	Q9PMG3	Q9PMG3 campylobact
835	5	17.9	371	16	Q981V3	Q981V3 rhizobium l	908	5	17.9	402	16	Q9PGC5	Q9PGC5 xylella fas
836	5	17.9	372	16	Q92TES	Q92TES rhizobium m	909	5	17.9	403	4	Q96CX6	Q96CX6 homo sapien
837	5	17.9	373	10	Q9SFC0	Q9SFC0 arabisdopsis	910	5	17.9	403	17	Q29552	Q29552 archaeoglob
838	5	17.9	373	10	Q94HI3	Q94HI3 oryza sativ	911	5	17.9	403	17	Q97C48	Q97C48 thermoplasm
839	5	17.9	373	17	Q29983	Q29983 archaeoglob	912	5	17.9	404	2	Q70029	Q70029 streptomyce
840	5	17.9	374	4	Q9H688	Q9H688 homo sapien	913	5	17.9	404	5	Q9VJE3	Q9VJE3 drosophila
841	5	17.9	374	10	Q9LHW8	Q9LHW8 oryza sativ	914	5	17.9	404	17	Q978N3	Q978N3 thermoplasm
842	5	17.9	375	2	Q9S214	Q9S214 streptomyce	915	5	17.9	405	2	Q54432	Q54432 lactococcus
843	5	17.9	375	2	Q59965	Q59965 synecococc	916	5	17.9	406	13	Q91619	Q91619 xenopus lae
844	5	17.9	375	13	Q9DEZ1	Q9DEZ1 xenopus lae	917	5	17.9	406	17	Q59601	Q59601 pyrococcus
845	5	17.9	376	10	Q9C955	Q9C955 arabisdopsis	918	5	17.9	406	17	Q9UXZ5	Q9UXZ5 pyrococcus
846	5	17.9	376	16	Q92WN1	Q92WN1 rhizobium m	919	5	17.9	408	2	Q924V4	Q924V4 streptomyce
847	5	17.9	377	4	Q9UQ19	Q9UQ19 homo sapien	920	5	17.9	408	2	Q48658	Q48658 lactococcus
848	5	17.9	377	5	Q21995	Q21995 caenorhabdi	921	5	17.9	408	12	Q9PYN7	Q9PYN7 xestia c-ni
849	5	17.9	377	12	Q9DUT6	Q9DUT6 porcine rub	922	5	17.9	408	16	Q9PJ96	Q9PJ96 campylobact
850	5	17.9	378	4	Q9X6Y6	Q9X6Y6 homo sapien	923	5	17.9	408	16	Q9CDQ3	Q9CDQ3 lactococcus
851	5	17.9	378	4	Q94920	Q94920 homo sapien	924	5	17.9	409	4	Q9NRV3	Q9NRV3 homo sapien
852	5	17.9	379	10	Q9LDO6	Q9LDO6 oryza sativ	925	5	17.9	409	5	Q9VKT4	Q9VKT4 drosophila
853	5	17.9	379	10	Q49316	Q49316 arabisdopsis	926	5	17.9	409	10	Q9SW87	Q9SW87 arachis hyp
854	5	17.9	379	16	Q92PD5	Q92PD5 rhizobium m	927	5	17.9	410	5	Q9TVZ6	Q9TVZ6 caenorhabdi
855	5	17.9	379	17	Q92PD5	Q92PD5 rhizobium m	928	5	17.9	411	2	Q9XDE0	Q9XDE0 bruceella me
856	5	17.9	380	10	Q9XEZ0	Q9XEZ0 nicotiana t	929	5	17.9	411	4	Q96J89	Q96J89 homo sapien
857	5	17.9	381	12	Q41941	Q41941 murid herpe	930	5	17.9	412	16	Q92PJ0	Q92PJ0 rhizobium m
858	5	17.9	383	4	Q9NXT9	Q9NXT9 homo sapien	931	5	17.9	413	5	Q25004	Q25004 homarus ame
859	5	17.9	383	10	Q23578	Q23578 arabisdopsis	932	5	17.9	413	6	Q9TU17	Q9TU17 ovis aries
860	5	17.9	383	11	Q91WC7	Q91WC7 mus musculu	933	5	17.9	413	16	Q9KLL5	Q9KLL5 vibrio chol
861	5	17.9	383	11	Q91W88	Q91W88 mus musculu	934	5	17.9	414	3	Q06674	Q06674 saccharomyc
862	5	17.9	383	16	Q55666	Q55666 synecocyst	935	5	17.9	414	4	Q92552	Q92552 homo sapien
863	5	17.9	384	16	Q9KPM7	Q9KPM7 vibrio chol	936	5	17.9	414	5	Q24892	Q24892 eimeria max
864	5	17.9	384	16	Q98Q07	Q98Q07 mycoplasma	937	5	17.9	416	2	Q9RJJ6	Q9RJJ6 streptomyce
865	5	17.9	385	16	Q92XM8	Q92XM8 rhizobium m	938	5	17.9	417	5	Q9BJW5	Q9BJW5 caenorhabdi
866	5	17.9	386	4	Q9H767	Q9H767 homo sapien	939	5	17.9	417	16	Q92PK5	Q92PK5 rhizobium m
867	5	17.9	386	5	Q62003	Q62003 branchiosto	940	5	17.9	419	5	Q62312	Q62312 caenorhabdi
868	5	17.9	387	4	Q9HAY8	Q9HAY8 homo sapien	941	5	17.9	419	12	Q910L6	Q910L6 white spot
869	5	17.9	387	4	Q96FW9	Q96FW9 homo sapien	942	5	17.9	419	17	Q9HN77	Q9HN77 halobacteri
870	5	17.9	387	5	Q46006	Q46006 caenorhabdi	943	5	17.9	420	2	Q9ZHQ1	Q9ZHQ1 streptomyce
871	5	17.9	387	5	Q9GVV5	Q9GVV5 drosophila	944	5	17.9	420	2	Q9RJQ6	Q9RJQ6 streptomyce
872	5	17.9	387	8	Q78342	Q78342 eimeria ten	945	5	17.9	420	2	Q93NC9	Q93NC9 myxococcus
873	5	17.9	389	5	Q95U18	Q95U18 drosophila	946	5	17.9	420	3	P78765	P78765 schizosacch
874	5	17.9	389	11	Q9EQ18	Q9EQ18 mus musculu	947	5	17.9	420	3	P78765	P78765 schizosacch
875	5	17.9	389	11	Q9ERJ9	Q9ERJ9 mus musculu	948	5	17.9	420	4	Q9H533	Q9H533 homo sapien
876	5	17.9	390	11	Q55223	Q55223 mus musculu	949	5	17.9	420	11	O88370	O88370 rattus norv
877	5	17.9	390	11	Q9D826	Q9D826 mus musculu	950	5	17.9	420	17	Q97BC2	Q97BC2 thermoplasm
878	5	17.9	390	16	Q9P126	Q9P126 campylobact	951	5	17.9	421	2	O68243	O68243 pantoea cit
879	5	17.9	390	16	P96446	P96446 rhizobium m	952	5	17.9	421	3	O74595	O74595 schizosacch
880	5	17.9	392	3	Q9FP65	Q9FP65 nigrospora	953	5	17.9	421	3	P78580	P78580 aspergillus
881	5	17.9	392	16	Q9RW09	Q9RW09 deinococcus	954	5	17.9	421	11	Q91XU3	Q91XU3 mus musculu
882	5	17.9	393	3	Q9USJ7	Q9USJ7 schizosacch	955	5	17.9	425	10	P93744	P93744 arabisdopsis
883	5	17.9	393	10	Q94DD4	Q94DD4 oryza sativ	956	5	17.9	425	16	P74565	P74565 synecocyst
884	5	17.9	394	16	Q9RRB5	Q9RRB5 deinococcus	957	5	17.9	425	16	Q98H12	Q98H12 rhizobium l
885	5	17.9	394	17	Q9UY77	Q9UY77 pyrococcus	958	5	17.9	426	13	Q9DGM2	Q9DGM2 brachydantio
886	5	17.9	395	2	Q9X411	Q9X411 chlamyidophi	959	5	17.9	428	10	Q9AQV2	Q9AQV2 oryza sativ
887	5	17.9	395	2	Q9X412	Q9X412 chlamyidophi	960	5	17.9	428	16	Q98M85	Q98M85 rhizobium l
888	5	17.9	395	2	Q9S6B3	Q9S6B3 chlamydia p	961	5	17.9	429	10	Q9T023	Q9T023 arabisdopsis
889	5	17.9	395	16	Q98L29	Q98L29 rhizobium l	962	5	17.9	430	2	Q9F2K3	Q9F2K3 streptomyce
890	5	17.9	396	10	Q9SEK6	Q9SEK6 persea amer	963	5	17.9	431	11	Q9Z304	Q9Z304 mus musculu
891	5	17.9	397	3	Q01972	Q01972 penicillium	964	5	17.9	431	16	Q9KCA6	Q9KCA6 bacillus ha
892	5	17.9	397	2	Q9ZHQ7	Q9ZHQ7 streptomyce	965	5	17.9	432	2	P77540	P77540 escherichia

966	5	17.9	432	2	Q9KYF7	Q9kyf7 streptomyce	1039	5	17.9	450	2	Q9F235	Q9f235 arthrobacte
967	5	17.9	433	1	Q93645	Q93645 pyrococcus	1040	5	17.9	451	2	Q9ADG6	Q9adg6 streptomyce
968	5	17.9	434	2	Q9AGL8	Q9agl8 arthrobacte	1041	5	17.9	451	4	Q9BSI4	Q9bsi4 homo sapien
969	5	17.9	435	16	P73403	P73403 synechocyst	1042	5	17.9	451	11	Q99JX3	Q99jx3 mus musculu
970	5	17.9	436	9	Q9JTM5	Q9jtm5 bacterioph	1043	5	17.9	451	16	Q9KFS9	Q9kfs9 vibrio chol
971	5	17.9	437	3	Q9P4U8	Q9p4u8 alternaria	1044	5	17.9	452	12	Q9YKI6	Q9yki6 hepatitis c
972	5	17.9	438	11	Q9CTA2	Q9cta2 mus musculu	1045	5	17.9	453	3	Q96VT0	Q96vt0 agrocyste pe
973	5	17.9	439	16	P73669	P73669 synechocyst	1046	5	17.9	453	16	Q9CDY7	Q9cdy7 lactococcus
974	5	17.9	440	2	Q9FA01	Q9fa01 arthrobacte	1047	5	17.9	454	5	Q9VZ14	Q9vz14 drosophila
975	5	17.9	441	2	Q9S4D6	Q9s4d6 streptomyce	1048	5	17.9	454	11	Q9R064	Q9r064 rattus norv
976	5	17.9	442	16	Q98RQ2	Q98fq2 rhizobium l	1049	5	17.9	454	16	Q9HT22	Q9ht22 pseudomonas
977	5	17.9	443	2	Q98305	Q98305 arthrobacte	1050	5	17.9	455	10	Q9ZVY5	Q9zvy5 arabidopsis
978	5	17.9	444	2	Q94823	Q94823 streptomyce	1051	5	17.9	457	10	Q9SU82	Q9su82 arabidopsis
979	5	17.9	445	2	Q9EZJ8	Q9ezj8 thermus aqu	1052	5	17.9	458	5	Q9GTU2	Q9gtu2 drosophila
980	5	17.9	446	11	Q9DC79	Q9dc79 mus musculu	1053	5	17.9	458	17	Q95215	Q95215 pyrococcus
981	5	17.9	447	11	Q9JMF5	Q9jmf5 mus musculu	1054	5	17.9	458	17	Q9YV129	Q9yv129 pyrococcus
982	5	17.9	448	2	Q9JPI6	Q9jpi6 neisseria m	1055	5	17.9	459	4	Q9Y330	Q9y330 homo sapien
983	5	17.9	449	2	Q9JPH9	Q9jph9 neisseria m	1056	5	17.9	459	10	P93816	P93816 arabidopsis
984	5	17.9	450	2	Q9JPR6	Q9jpr6 neisseria m	1057	5	17.9	459	11	Q9Z150	Q9z150 mus musculu
985	5	17.9	451	2	Q9JPR5	Q9jpr5 neisseria m	1058	5	17.9	459	16	Q9KBG2	Q9kbg2 bacillus ha
986	5	17.9	452	2	Q9JPR4	Q9jpr4 neisseria m	1059	5	17.9	460	10	Q94G86	Q94g86 olea europ
987	5	17.9	453	2	Q9JPR3	Q9jpr3 neisseria m	1060	5	17.9	461	2	Q9KI52	Q9ki52 enterococu
988	5	17.9	454	2	Q9JPR2	Q9jpr2 neisseria m	1061	5	17.9	461	16	Q987A9	Q987a9 rhizobium l
989	5	17.9	455	2	Q9JPR1	Q9jpr1 neisseria m	1062	5	17.9	463	4	Q92981	Q92981 homo sapien
990	5	17.9	456	2	Q9JPR0	Q9jpr0 neisseria m	1063	5	17.9	463	4	Q976019	Q976019 homo sapien
991	5	17.9	457	2	Q9JPP9	Q9jpp9 neisseria m	1064	5	17.9	464	4	Q9BT22	Q9bt22 homo sapien
992	5	17.9	458	2	Q9JPP8	Q9jpp8 neisseria m	1065	5	17.9	464	4	Q9P2Y2	Q9p2y2 homo sapien
993	5	17.9	459	2	Q9JPP7	Q9jpp7 neisseria m	1066	5	17.9	465	11	Q99ND8	Q99nd8 rattus norv
994	5	17.9	460	2	Q9JPP6	Q9jpp6 neisseria m	1067	5	17.9	465	16	Q9Z1X6	Q9z1x6 helicobacte
995	5	17.9	461	2	Q9JPP5	Q9jpp5 neisseria m	1068	5	17.9	467	16	Q97N88	Q97n88 streptococ
996	5	17.9	462	2	Q9JPP4	Q9jpp4 neisseria m	1069	5	17.9	469	16	Q97452	Q97452 aquifex ae
997	5	17.9	463	2	Q9JPP3	Q9jpp3 neisseria m	1070	5	17.9	469	17	Q97CF0	Q97cf0 thermoplas
998	5	17.9	464	2	Q9L6H3	Q9l6h3 neisseria g	1071	5	17.9	470	3	Q9USW2	Q9usw2 schizosac
999	5	17.9	465	2	Q9L6H2	Q9l6h2 neisseria g	1072	5	17.9	470	3	Q9FN78	Q9fn78 arabidopsis
1000	5	17.9	466	2	Q9L6H1	Q9l6h1 neisseria g	1073	5	17.9	472	12	Q91FX2	Q91fx2 chilo iride
1001	5	17.9	467	16	Q9PCP0	Q9pcp0 xyella fas	1074	5	17.9	472	16	Q981V5	Q981v5 rhizobium l
1002	5	17.9	468	16	Q9JRK38	Q9jrk38 neisseria m	1075	5	17.9	473	16	Q92SF1	Q92sf1 rhizobium m
1003	5	17.9	469	16	Q9JQX7	Q9jqx7 neisseria m	1076	5	17.9	474	16	Q9KFD0	Q9kfd0 bacillus ha
1004	5	17.9	470	17	Q972P1	Q972p1 sulfolobus	1077	5	17.9	475	2	Q929P3	Q929p3 bacillus ha
1005	5	17.9	471	4	Q9B2D5	Q9bzd5 homo sapien	1078	5	17.9	475	3	Q94315	Q94315 schizosac
1006	5	17.9	472	4	Q9BX97	Q9bx97 homo sapien	1079	5	17.9	475	3	Q96092	Q96092 halocynthia
1007	5	17.9	473	4	Q9H0M8	Q9h0m8 homo sapien	1080	5	17.9	475	10	Q9FWN3	Q9fnw3 oryza sativ
1008	5	17.9	474	11	Q9CX49	Q9cxy49 mus musculu	1081	5	17.9	476	4	Q9UFK0	Q9ufk0 homo sapien
1009	5	17.9	475	16	Q98BZ0	Q98dz0 rhizobium l	1082	5	17.9	476	17	Q928858	Q928858 archaeoglob
1010	5	17.9	476	4	Q96DH4	Q96dh4 homo sapien	1083	5	17.9	478	16	P76103	P76103 escherichia
1011	5	17.9	477	4	Q96D77	Q96d77 homo sapien	1084	5	17.9	479	2	Q93JF0	Q93jf0 streptomyce
1012	5	17.9	478	5	Q9TZ81	Q9tzb1 caenorhabdi	1085	5	17.9	479	4	Q96MA6	Q96ma6 homo sapien
1013	5	17.9	479	16	Q910I2	Q910i2 pseudomonas	1086	5	17.9	479	16	Q9PJ47	Q9pj47 campylobact
1014	5	17.9	480	2	Q9X3V9	Q9x3v9 zymomonas m	1087	5	17.9	480	10	Q9LSX4	Q9lsx4 arabidopsis
1015	5	17.9	481	2	Q93H63	Q93h63 streptomyce	1088	5	17.9	480	16	Q925231	Q925231 helicobacte
1016	5	17.9	482	10	Q9AX12	Q9axi2 euglena gra	1089	5	17.9	482	10	Q9FVW4	Q9fvw4 arabidopsis
1017	5	17.9	483	16	Q91455	Q91455 bacillus su	1090	5	17.9	484	3	Q9Y7U9	Q9y7u9 schizosac
1018	5	17.9	484	5	Q9XTB7	Q9xtb7 caenorhabdi	1091	5	17.9	484	16	Q9RZ17	Q9rcz17 deinococcus
1019	5	17.9	485	12	Q98168	Q98168 hepatitis c	1092	5	17.9	484	16	Q98CS0	Q98cs0 rhizobium l
1020	5	17.9	486	12	Q98169	Q98169 hepatitis c	1093	5	17.9	485	10	Q9FTS8	Q9fts8 oryza sativ
1021	5	17.9	487	12	Q9ENM4	Q9enw4 hepatitis c	1094	5	17.9	485	13	Q9PWC0	Q9pwc0 gallus gall
1022	5	17.9	488	12	Q9ENW3	Q9enw3 hepatitis c	1095	5	17.9	486	5	Q9W3M9	Q9w3m9 drosophila
1023	5	17.9	489	12	Q9ENW2	Q9enw2 hepatitis c	1096	5	17.9	486	15	Q90634	Q90634 human immun
1024	5	17.9	490	12	Q98481	Q98481 hepatitis c	1097	5	17.9	486	16	Q97Q12	Q97q12 streptococ
1025	5	17.9	491	12	Q98482	Q98482 hepatitis c	1098	5	17.9	488	16	Q9CL83	Q9cl83 pasteurella
1026	5	17.9	492	12	Q98484	Q98484 hepatitis c	1099	5	17.9	489	5	Q9V9U3	Q9v9u3 drosophila
1027	5	17.9	493	12	Q98485	Q98485 hepatitis c	1100	5	17.9	490	16	P95098	P95098 mycobacteri
1028	5	17.9	494	12	Q98489	Q98489 hepatitis c	1101	5	17.9	491	4	Q9NUR0	Q9nur0 homo sapien
1029	5	17.9	495	12	Q98490	Q98490 hepatitis c	1102	5	17.9	491	4	Q9BR58	Q9brs8 homo sapien
1030	5	17.9	496	12	Q9YK96	Q9yk96 hepatitis c	1103	5	17.9	492	6	Q9N2C7	Q9n2c7 canis famil
1031	5	17.9	497	12	Q98493	Q98493 hepatitis c	1104	5	17.9	492	11	Q9D3J0	Q9d3j0 mus musculu
1032	5	17.9	498	12	Q98494	Q98494 hepatitis c	1105	5	17.9	494	5	Q94275	Q94275 trypanosoma
1033	5	17.9	499	12	Q98495	Q98495 hepatitis c	1106	5	17.9	494	17	Q970C5	Q970c5 sulfolobus
1034	5	17.9	500	12	Q98496	Q98496 hepatitis c	1107	5	17.9	495	2	Q9JP38	Q9jp38 pseudomonas
1035	5	17.9	501	12	Q98497	Q98497 hepatitis c	1108	5	17.9	496	3	Q94468	Q94468 schizosac
1036	5	17.9	502	16	Q9HP62	Q9ht62 pseudomonas	1109	5	17.9	496	17	Q926819	Q926819 methanother
1037	5	17.9	503	10	Q9FEV0	Q9fev0 picea abies	1110	5	17.9	496	17	Q97V18	Q97v18 sulfolobus
1038	5	17.9	504	10	Q9LXN1	Q9lxl1 arabidopsis	1111	5	17.9	497	2	Q9KGQ6	Q9kgq6 vibrio para

1112	5	17.9	497	9	Q9MCC5	Q9mcc5 bacterioph	1185	5	17.9	554	16	Q92TT6	Q92tt6 rhizobium m
1113	5	17.9	498	2	Q47385	Q47385 erwinia car	1186	5	17.9	557	2	Q9AIS7	Q9ais7 chlamydomphi
1114	5	17.9	498	10	Q9XIM8	Q9xim8 arabidopsis	1187	5	17.9	557	10	Q9M1E1	Q9mie1 arabidopsis
1115	5	17.9	499	2	Q52115	Q52115 erwinia ste	1188	5	17.9	558	2	P94664	P94664 chlamydomphi
1116	5	17.9	499	10	Q9FX04	Q9fx04 arabidopsis	1189	5	17.9	560	10	Q43319	Q43319 artemisia a
1117	5	17.9	500	4	Q99526	Q99526 homo sapien	1190	5	17.9	561	8	Q34177	Q34177 cepaea nemo
1118	5	17.9	501	16	Q55762	Q55762 synecocyst	1191	5	17.9	562	16	Q9HYU3	Q9hyu3 pseudomonas
1119	5	17.9	501	17	Q9UYD7	Q9uyd7 pyrococcus	1192	5	17.9	563	12	Q9Q8G8	Q9q8g8 myxoma viru
1120	5	17.9	502	5	Q61495	Q61495 drosophila	1193	5	17.9	564	2	O24852	O24852 aeromonas p
1121	5	17.9	502	5	Q9VUD7	Q9vud7 drosophila	1194	5	17.9	564	12	O72736	O72736 cowpox viru
1122	5	17.9	506	5	Q18501	Q18501 caenorhabdi	1195	5	17.9	564	16	Q9I2B7	Q9i2b7 pseudomonas
1123	5	17.9	506	17	Q30295	Q30295 archaeoglob	1196	5	17.9	565	10	Q9M292	Q9m292 arabidopsis
1124	5	17.9	508	11	Q923B6	Q923b6 mus musculu	1197	5	17.9	565	16	P73143	P73143 synecocyst
1125	5	17.9	509	2	Q9KZV2	Q9kzv2 streptomyce	1198	5	17.9	566	3	Q9P522	Q9p522 neurospora
1126	5	17.9	509	8	Q9BBG5	Q9bbg5 hydrobryum	1199	5	17.9	566	10	Q43318	Q43318 artemisia a
1127	5	17.9	509	8	Q9BBG4	Q9bbg4 hydrobryum	1200	5	17.9	567	16	Q9A929	Q9a929 caulobacter
1128	5	17.9	510	4	Q96NY8	Q96ny8 homo sapien							
1129	5	17.9	510	4	Q96K15	Q96k15 homo sapien							
1130	5	17.9	510	11	Q9DCV6	Q9dcv6 mus musculu							
1131	5	17.9	512	5	Q96091	Q96091 halocynthia							
1132	5	17.9	513	11	Q63184	Q63184 rattus norv							
1133	5	17.9	514	4	Q60381	Q60381 homo sapien							
1134	5	17.9	514	4	O14790	O14790 homo sapien							
1135	5	17.9	514	4	Q96AJ2	Q96aj2 homo sapien							
1136	5	17.9	515	4	Q96EU9	Q96eu9 homo sapien							
1137	5	17.9	515	4	Q969R4	Q969r4 homo sapien							
1138	5	17.9	515	17	Q58210	Q58210 pyrococcus							
1139	5	17.9	516	12	Q99711	Q99711 bovine aden							
1140	5	17.9	517	4	Q96GD8	Q96gd8 homo sapien							
1141	5	17.9	517	16	Q98BX2	Q98bx2 rhizobium 1							
1142	5	17.9	518	4	Q9NP81	Q9np81 homo sapien							
1143	5	17.9	518	4	Q9BUC3	Q9buc3 homo sapien							
1144	5	17.9	518	4	Q9NVX7	Q9nvx7 homo sapien							
1145	5	17.9	518	4	Q9NV76	Q9nv76 homo sapien							
1146	5	17.9	518	11	Q9CWF1	Q9cwf1 mus musculu							
1147	5	17.9	518	11	Q9JUL8	Q9jyl8 mus musculu							
1148	5	17.9	518	16	Q9RVW8	Q9rvw8 deinococcus							
1149	5	17.9	519	2	Q9LBT7	Q9lbt7 microcystis							
1150	5	17.9	520	10	Q9FZL2	Q9fzl2 arabidopsis							
1151	5	17.9	520	12	O41104	O41104 paramecium							
1152	5	17.9	521	15	Q9WPE6	Q9wpp6 chimpanzee							
1153	5	17.9	522	4	Q96GR3	Q96gr3 homo sapien							
1154	5	17.9	523	17	Q97W17	Q97w17 sulfolobus							
1155	5	17.9	524	4	Q96IL3	Q96il3 homo sapien							
1156	5	17.9	525	10	Q9FX10	Q9fx10 arabidopsis							
1157	5	17.9	526	16	Q53278	Q53278 mycobacteri							
1158	5	17.9	526	16	Q9HZ11	Q9hzi1 pseudomonas							
1159	5	17.9	527	2	Q93Q40	Q93q40 myxococcus							
1160	5	17.9	527	12	Q99GX7	Q99gy7 helicoverp							
1161	5	17.9	528	2	P70733	P70733 aeromonas p							
1162	5	17.9	529	2	Q9K4D5	Q9k4d5 streptomyce							
1163	5	17.9	529	16	Q97PS2	Q97ps2 streptococc							
1164	5	17.9	530	4	Q96IU9	Q96iu9 homo sapien							
1165	5	17.9	533	5	Q9B147	Q9b147 caenorhabdi							
1166	5	17.9	534	2	Q9ZAW1	Q9zaw1 chlamydomphi							
1167	5	17.9	534	2	Q9ZAW0	Q9zaw0 chlamydomphi							
1168	5	17.9	535	5	Q961W5	Q961w5 drosophila							
1169	5	17.9	537	16	Q05459	Q05459 mycobacteri							
1170	5	17.9	539	11	Q9QX32	Q9qy32 mus musculu							
1171	5	17.9	541	5	Q9V162	Q9v162 drosophila							
1172	5	17.9	541	5	O15681	O15681 toxoplasma							
1173	5	17.9	541	10	Q22887	Q22887 arabidopsis							
1174	5	17.9	541	10	Q9ASL7	Q9asl7 oryza sativ							
1175	5	17.9	543	6	Q9N010	Q9n010 macaca fasc							
1176	5	17.9	544	2	Q9X8Z2	Q9x8z2 streptomyce							
1177	5	17.9	544	16	Q9A195	Q9a195 streptococc							
1178	5	17.9	545	10	Q41728	Q41728 zea mays (m							
1179	5	17.9	545	16	Q91165	Q91165 pseudomonas							
1180	5	17.9	547	16	Q9A9A8	Q9a9a8 caulobacter							
1181	5	17.9	548	16	O05318	O05318 mycobacteri							
1182	5	17.9	551	4	Q9F613	Q9f613 homo sapien							
1183	5	17.9	553	3	Q9U0B3	Q9uub3 schizosacch							
1184	5	17.9	554	11	Q62835	Q62835 rattus norv							

RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RL halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001520; BAB07733.1; -.
DR InterPro; IPR001440; TPR.
KW Complete proteome.
SQ SEQUENCE 370 AA; 43899 MW; 5163D42663A9FAFD CRC64;

Query Match 25.0%; Score 7; DB 16; Length 370;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KENLPSL 16
| | | | |
Db 357 KENLPSL 363

RESULT 3
Q91MY7 PRELIMINARY; PRT; 569 AA.
ID Q91MY7
AC Q91MY7
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE LSDV019 KELCH-LIKE PROTEIN.
GN LSDV019
OS lumpy skin disease virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Capripoxvirus.
OX NCBI_TaxID=59509;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEETHLING 2490;
RX MEDLINE=21329495; PubMed=11435593;
RA Tulman E.R., Alfonso C.L., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT "Genome of lumpy skin disease virus.";
RL J. Virol. 75:7122-7130(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NEETHLING 2490;
RA Tulman E.R., Alfonso C.L., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF325528; AAK84980.1; -.
SQ SEQUENCE 569 AA; 66876 MW; 5034D8170A348FDB CRC64;

Query Match 25.0%; Score 7; DB 12; Length 569;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KENLPSL 16
| | | | |
Db 368 KENLPSL 374

RESULT 4
Q9Y5D0 PRELIMINARY; PRT; 817 AA.
ID Q9Y5D0
AC Q9Y5D0
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PROTOCADHERIN GAMMA A7 SHORT FORM PROTEIN.
GN PCDH-GAMMA-A7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99308636; PubMed=10380929;

RA Wu Q., Maniatis T.;
RT "A striking organization of a large family of human neural cadherin-
RT like cell adhesion genes.";
RL Cell 97:779-790(1999).
CC -!- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.
DR EMBL; AF152514; AAD43774.1; -.
DR HSSP; P15116; INCJ.
DR InterPro; IPR002126; Cadherin.
DR Pfam; PF00028; cadherin; 6.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 6.
DR PROSITE; PS00232; CADHERIN_1; 1.
DR PROSITE; PS00268; CADHERIN_2; 6.
KW Calcium-binding; Cell adhesion; Glycoprotein.
SQ SEQUENCE 817 AA; 89717 MW; 4B956F3AFB34F022 CRC64;

Query Match 25.0%; Score 7; DB 4; Length 817;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 CKENLPS 15
| | | | |
Db 800 CKENLPS 806

RESULT 5
Q9Y5G6 PRELIMINARY; PRT; 932 AA.
ID Q9Y5G6
AC Q9Y5G6
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PROTOCADHERIN GAMMA A7.
GN PCDH-GAMMA-A7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=99308636; PubMed=10380929;
RA Wu Q., Maniatis T.;
RT "A striking organization of a large family of human neural cadherin-
RT like cell adhesion genes.";
RL Cell 97:779-790(1999).
CC -!- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.
DR EMBL; AF152327; AAD43721.1; -.
DR HSSP; P15116; INCJ.
DR InterPro; IPR002126; Cadherin.
DR Pfam; PF00028; cadherin; 6.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 6.
DR PROSITE; PS00232; CADHERIN_1; 1.
DR PROSITE; PS00268; CADHERIN_2; 6.
KW Calcium-binding; Cell adhesion; Glycoprotein.
SQ SEQUENCE 932 AA; 101721 MW; 9614AE69C0CA2371 CRC64;

Query Match 25.0%; Score 7; DB 4; Length 932;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 CKENLPS 15
| | | | |
Db 800 CKENLPS 806

RESULT 6
O87046 PRELIMINARY; PRT; 94 AA.
ID O87046
AC O87046;

DT 01-NOV-1998 (TReMBLrel. 08, Created)
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
 DE 244F PROTEIN.
 GN 244F.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=217561;
 RA Fallarino A.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ231102; CAAL3144.1; -.
 SQ SEQUENCE 94 AA; 10029 MW; 34F10CDF8E109262 CRC64;

Query Match 21.4%; Score 6; DB 2; Length 94;

Best Local Similarity 100.0%; Pred. No. 45;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KENLPS 15

Db 83 KENLPS 88

RESULT 7

ID Q53981 PRELIMINARY; PRT; 103 AA.
 AC Q53981;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE HYPOTHETICAL 11.4 KDA PROTEIN (FRAGMENT).
 OS Shigella dysenteriae.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Shigella.
 OX NCBI_TaxID=622;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SH16;
 RX MEDLINE=88062685; PubMed=2824781;
 RA Matsutani S., Ohtsubo H., Maeda Y., Ohtsubo E.;
 RT "Isolation and characterization of IS elements repeated in the
 RT bacterial chromosome."
 RL J. Mol. Biol. 196:445-455(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SH16;
 RX MEDLINE=97346058; PubMed=9202481;
 RA Tomimaga A.;
 RT "The site-specific recombinase encoded by pind in Shigella dysenteriae
 RT is due to the presence of a defective Mu prophage."
 RL Microbiology 143:2057-2063(1997).
 DR EMBL; D63765; BAA09847.1; -.
 DR InterPro; IPR000825; UPF0051.
 DR Pfam; PF01458; UPF0051; 1.
 KW Hypothetical protein.
 FT NON_TER 1
 FT NON_TER 103
 SQ SEQUENCE 103 AA; 11414 MW; C73C6E428074CE4B CRC64;

Query Match 21.4%; Score 6; DB 2; Length 103;

Best Local Similarity 100.0%; Pred. No. 49;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6

Db 93 TGSALT 98

RESULT 8

Q9RNC6
 ID Q9RNC6 PRELIMINARY; PRT; 104 AA.
 AC Q9RNC6;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE ALPHA KETOGLUTARATE DEPENDENT DIOXYGENASE (FRAGMENT).
 GN TFDA.
 OS Alcaligenes sp. CSI.
 OC Plasmid pCS1.
 OG Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
 OC Alcaligenes.
 OX NCBI_TaxID=104732;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MCPT CS;
 RA Snejkal C.W., Vallaeys T., Burton S.K., Lappin-Scott H.M.;
 RT "Isolation and partial sequencing of tfda gene from a mecoprop
 RT degrading Alcaligenes species."
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF181982; AAF62390.2; -.
 KW Dioxygenase; Plasmid.
 FT NON_TER 1
 FT NON_TER 104
 SQ SEQUENCE 104 AA; 11856 MW; A02C8DBF00A2CDD3 CRC64;

Query Match 21.4%; Score 6; DB 2; Length 104;

Best Local Similarity 100.0%; Pred. No. 49;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25

Db 11 QGLRAE 16

RESULT 9

ID Q9RAF1 PRELIMINARY; PRT; 104 AA.
 AC Q9RAF1;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE ALPHA KETOGLUTARATE DEPENDENT DIOXYGENASE (FRAGMENT).
 GN TFDA
 OS Ralstonia sp. CS2.
 OG Plasmid pCS2.
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 OC Ralstonia.
 OX NCBI_TaxID=115801;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS2;
 RA Snejkal C.W., Vallaeys T., Burton S.K., Lappin-Scott H.M.;
 RT "Isolation and partial sequencing of tfda gene from a mecoprop
 RT degrading Ralstonia species."
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF182758; AAF02720.2; -.
 KW Dioxygenase; Plasmid.
 FT NON_TER 1
 FT NON_TER 104
 SQ SEQUENCE 104 AA; 11856 MW; A02C8DBF00A2CDD3 CRC64;

Query Match 21.4%; Score 6; DB 2; Length 104;

Best Local Similarity 100.0%; Pred. No. 49;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25

Db 11 QGLRAE 16

```

RESULT 10
Q47965
ID Q47965 PRELIMINARY; PRT; 105 AA.
AC Q47965;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 10, Last annotation update)
DE 2,4-DICHLOROPHENOXYACETATE DIOXYGENASE (FRAGMENT).
GN TFDA.
OS Halomonadaceae gen. sp. strain I-18.
OC Bacteria; Proteobacteria; gamma subdivision; Halomonadaceae.
OX NCBI_TaxID=38770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-I-18;
RX MEDLINE=96254003; PubMed=8704953;
RA Maltseva O., McGowan C., Fulthorpe R., Oriol P.;
RT "Degradation of 2,4-dichlorophenoxyacetic acid by haloalkaliphilic bacteria.";
RL Microbiology 142:1115-1122(1996).
DR EMBL; U22499; AAB18140.1; -.
KW Dioxxygenase.
FT NON_TER 1 105
FT NON_TER 105 105
SQ SEQUENCE 105 AA; 12025 MW; 09CB5BAB51E788D4 CRC64;

Query Match 21.4%; Score 6; DB 2; Length 105;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
Db 12 QGLRAE 17

RESULT 11
P74954
ID P74954 PRELIMINARY; PRT; 106 AA.
AC P74954;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ALPHA KETOGLUTARATE DIOXYGENASE (FRAGMENT).
GN TFDA.
OS Variovorax paradoxus.
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;
OC Variovorax.
OX NCBI_TaxID=34073;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TV1;
RA Vallaeys T., Fulthorpe R.R., Wright A.M., Soulas G., Weightman A.J.;
RT "Isolation and partial sequencing of tfda gene from the 2,4-dichlorophenoxyacetic acid degrading bacterium Variovorax paradoxus TV1."
RT Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U65531; AAB08969.1; -.
KW Dioxxygenase.
FT NON_TER 1 106
FT NON_TER 106 106
SQ SEQUENCE 106 AA; 12135 MW; 15D31C4CCF9E6F13 CRC64;

Query Match 21.4%; Score 6; DB 2; Length 106;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
Db 13 QGLRAE 18

```

```

RESULT 12
Q9RAD3
ID Q9RAD3 PRELIMINARY; PRT; 109 AA.
AC Q9RAD3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE 2,4-D-ALPHA KETOGLUTARATE DIOXYGENASE (FRAGMENT).
GN TFDA.
OS Rhodoferax sp. P230.
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;
OC Rhodoferax.
OX NCBI_TaxID=103036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-P230;
RA Mueller R.H., Kleinsteinuber S., Babel W.;
RT "Physiological and genetic characteristics of two bacterial strains utilizing phenoxypyronate- and phenoxacetate-herbicides.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF176240; AAD55077.1; -.
KW Dioxxygenase.
FT NON_TER 1 109
FT NON_TER 109 109
SQ SEQUENCE 109 AA; 12433 MW; AECE1DF393A16F4 CRC64;

Query Match 21.4%; Score 6; DB 2; Length 109;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
Db 16 QGLRAE 21

```

```

RESULT 13
Q93T16
ID Q93T16 PRELIMINARY; PRT; 109 AA.
AC Q93T16;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE 2,4-D-ALPHA-KETOGLUTARATE DIOXYGENASE (FRAGMENT).
GN TFDA.
OS Comamonas acidovorans (Pseudomonas acidovorans).
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Delftia.
OX NCBI_TaxID=80866;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-P4A;
RA Hoffmann D., Kleinsteinuber S., Mueller R.H., Babel W.;
RT "Development and application of PCR primers for the detection of the tfd genes in Delftia acidovorans P4a involved in the degradation of 2,4-D."
RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF377325; AAK57006.1; -.
KW Dioxxygenase.
FT NON_TER 1 109
FT NON_TER 109 109
SQ SEQUENCE 109 AA; 12486 MW; 29EDE5DBE6C13237 CRC64;

Query Match 21.4%; Score 6; DB 2; Length 109;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
Db 16 QGLRAE 21

RESULT 14

```

Q985Y1
ID Q985Y1 PRELIMINARY; PRT; 133 AA.
AC Q985Y1
DT 01-OCT-2001 (TReMBLrel. 18, Created)
DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
DE MLL7476 PROTEIN.
GN MLL7476.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AF003003; BAB50955.1; -.
KW Complete proteome.
SQ SEQUENCE 133 AA; 14765 MW; 7E8FD066159E231A CRC64;

Query Match 21.4%; Score 6; DB 16; Length 133;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GLRAEE 26
Db 81 GLRAEE 86

RESULT 15
Q98EG4 PRELIMINARY; PRT; 141 AA.
ID Q98EG4
AC Q98EG4
DT 01-OCT-2001 (TReMBLrel. 18, Created)
DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
DE OSMOTICALLY INDUCIBLE PROTEIN C, PUTATIVE ATP/GTP BINDING
DE PROTEIN.
GN MLR4254.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AF003003; BAB50955.1; -.
DR InterPro; IPR003718; OsmC.
DR Pfam; PF02566; OsmC; 1.
KW Complete proteome.
SQ SEQUENCE 141 AA; 14490 MW; 323434B70CA58837 CRC64;

Query Match 21.4%; Score 6; DB 16; Length 141;
Best Local Similarity 100.0%; Pred. No. 65;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
Db 94 TGSALT 99

RESULT 16
Q9CKW0 PRELIMINARY; PRT; 142 AA.
ID Q9CKW0
AC Q9CKW0
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE ATPC.
GN ATPC OR PMI495.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FW70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; AE006185; AAK03579.1; -.
DR HSSP; P00832; IAQT.
DR InterPro; IPR001469; ATP-synt_DE.
DR Pfam; PF00401; ATP-synt_DE; 1.
DR ProDom; PD000944; ATP-synt_DE; 1.
KW Complete proteome.
SQ SEQUENCE 142 AA; 15492 MW; FF2EALAF0326E860 CRC64;

Query Match 21.4%; Score 6; DB 16; Length 142;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 RAENI 28
Db 100 RAENI 105

RESULT 17
Q96MK6 PRELIMINARY; PRT; 148 AA.
ID Q96MK6
AC Q96MK6
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE CDNA FLJ32216 FIS, CLONE PLACE6003745.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Ishibashi T., Kanohori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA Hotuta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J., Chiba Y.,
RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK056778; BAB71282.1; -.
SQ SEQUENCE 148 AA; 16668 MW; AFDCCF8784DA65B CRC64;

Query Match 21.4%; Score 6; DB 4; Length 148;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 ENLPSL 16
 |||||
 Db 47 ENLPSL 52

RESULT 18

Q92294 PRELIMINARY; PRT; 153 AA.
 AC Q92294;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DE HYPOTHETICAL PROTEIN SMA1100.
 GN SMA1100.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OG Bacteroidetes (megaplasmid 1).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396509; PubMed=11481432;
 RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
 RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
 RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
 RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
 RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
 RT "Nucleotide sequence and predicted functions of the entire
 Sinorhizobium meliloti pSymA megaplasmid.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
 DR EMBL; AF007249; AAK65256.1; -;
 KW Hypothetical protein; Plasmid; Complete proteome.
 SQ SEQUENCE 153 AA; 16367 MW; F25A68FC70C6D88F CRC64;

Query Match 21.4%; Score 6; DB 16; Length 153;
 Best Local Similarity 100.0%; Pred. No. 69;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 GLRAEE 26
 |||||
 Db 22 GLRAEE 27

RESULT 19

Q9MCD9 PRELIMINARY; PRT; 164 AA.
 AC Q9MCD9;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DE TAIL TUBE PROTEIN GP19.
 GN 19.
 OS Bacteriophage RB49.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
 OC T4-like phages.
 OX NCBI_TaxID=50948;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20566703; PubMed=11114936;
 RA Tetart F., Desplats C., Kutateladze M., Monod C., Ackermann H.W.,
 RA Krisch H.M.;
 RT "Phylogeny of the major head and tail genes of the wide-ranging T4-
 type bacteriophages.";
 RL J. Bacteriol. 183:358-366(2001).
 DR EMBL; AF223003; AAF65258.1; -;
 SQ SEQUENCE 164 AA; 18409 MW; A3AD15EE9BF00C0B CRC64;

Query Match 21.4%; Score 6; DB 9; Length 164;
 Best Local Similarity 100.0%; Pred. No. 74;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGSALT 6
 |||||
 Db 97 TGSALT 102

RESULT 20

Q9W0V0 PRELIMINARY; PRT; 166 AA.
 AC Q9W0V0;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DE CG6845 PROTEIN.
 GN CG6845.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003467; AAF47333.1; -;
 DR FlyBase; FBgn0035099; CG6845.
 SQ SEQUENCE 166 AA; 18333 MW; FA0B749AD43F7602 CRC64;

Query Match

21.4%; Score 6; DB 5; Length 166;

```
Best Local Similarity 100.0%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LPSLCS 18
Db 65 LPSLCS 70

RESULT 21
Q9KFH6
ID Q9KFH6 PRELIMINARY; PRT; 174 AA.
AC Q9KFH6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE BH0503 PROTEIN.
GN BH0503.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001508; BAB04222.1; -.
DR InterPro; IPR000182; Acetyltransf_GCN5.
DR Pfam; PF00583; Acetyltransf; 1.
KW Complete proteome.
SQ SEQUENCE 174 AA; 19690 MW; 11D88EB1C644F5E3 CRC64;

Query Match 21.4%; Score 6; DB 16; Length 174;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KENLPS 15
Db 138 KENLPS 143

RESULT 22
O30273
ID O30273 PRELIMINARY; PRT; 175 AA.
AC O30273;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE ACETYL-COA DECARBOXYLASE/SYNTHASE, SUBUNIT EPSILON (CDHB-2).
GN AF2398.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kervilavage A.R., Graham D.E., Kyriades N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidner J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
```

```
Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
DR EMBL; AE001110; AAB91265.1; -.
DR TIGR; AF2398; -.
DR InterPro; IPR003704; CO_dh.
DR Pfam; PF02552; CO_dh; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 175 AA; 19702 MW; 26CA5B99DB3135D5 CRC64;

Query Match 21.4%; Score 6; DB 17; Length 175;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
Db 68 TGSALT 73

RESULT 23
Q92NS5
ID Q92NS5 PRELIMINARY; PRT; 184 AA.
AC Q92NS5;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PROBABLE RNA POLYMERASE SIGMA FACTOR PROTEIN.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21368234; PubMed=11474104;
RA Gallibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F.,
RA Barloy-Hubier F., Barnett M.J., Becker A., Boistard P., Bothe G.,
RA Boutry M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P.,
RA Cowie A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F.,
RA Gloux S., Godrie T., Goffeau A., Golding B., Gouzy J., Gurjal M.,
RA Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,
RA Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Lelaure V.,
RA Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B.,
RA Ramsperger U., Surzycki R., Thebault P., Vandenbol M.,
RA Vorhoeelter F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;
RT "The composite genome of the legume symbiont Sinorhizobium meliloti.";
RL Science 293:668-672(2001).
DR EMBL; AL591789; CAC46681.1; -.
KW Complete proteome.
SQ SEQUENCE 184 AA; 21062 MW; 3F5DCC1B7D61610D CRC64;

Query Match 21.4%; Score 6; DB 16; Length 184;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GLRAEE 26
Db 102 GLRAEE 107

RESULT 24
P74447
ID P74447 PRELIMINARY; PRT; 186 AA.
AC P74447;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE HYPOTHETICAL 20.6 KDA PROTEIN.
GN SLR0148.
```

OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=11148;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hirose M., Suglura M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
 RA Shimpō S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 Synechocystis sp. strain PCC6803. II. Sequence determination of the
 entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 DR EMBL; D90915; BAA18548.1; -;
 DR InterPro; IPR001041; Ferredoxin.
 DR Pfam; PF00111; fer2; 1.
 KW Complete proteome; Hypothetical protein; Iron-sulfur.
 SQ SEQUENCE 186 AA; 20625 MW; 4B199CC912A8D9B3 CRC64;

Query Match 21.4%; Score 6; DB 16; Length 186;
 Best Local Similarity 100.0%; Pred. No. 83;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 RAEENI 28
 Db 140 RAEENI 145

RESULT 25

ID Q9HWL0 PRELIMINARY; PRT; 187 AA.
 AC Q9HWL0;
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE PROBABLE PROTEASE.
 GN PA4171.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 CC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 DR EMBL; AE004833; AAG07558.1; -;
 DR MEROPS; C56.UPW; -;
 DR InterPro; IPR002818; ThiJ.
 DR Pfam; PF01965; ThiJ; 1.
 KW Protease; Complete proteome.
 SQ SEQUENCE 187 AA; 20076 MW; F8512BC780E29BE5 CRC64;

Query Match 21.4%; Score 6; DB 16; Length 187;
 Best Local Similarity 100.0%; Pred. No. 83;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 QGLRAE 25
 Db 68 QGLRAE 73

RESULT 26
 Q99RNA

ID Q99RNA PRELIMINARY; PRT; 191 AA.
 AC Q99RNA;
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE SA2183 PROTEIN (HYPOTHETICAL PROTEIN SAV2395).
 GN SA2183 OR SAV2395.
 OS Staphylococcus aureus (strain N315), and
 Staphylococcus aureus (strain Mu50).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 CC Bacillus/Staphylococcus group; Staphylococcus.
 OX NCBI_TaxID=158879; 158878;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.aureus (strain N315), and S.aureus (strain Mu50);
 RX MEDLINE=21311952; PubMed=11418146;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
 RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus
 aureus.";
 RL Lancet 357:1225-1240(2001).
 DR EMBL; AP003137; BAB43485.1; -;
 DR EMBL; AP003365; BAB58557.1; -;
 DR InterPro; IPR003765; Nitrate_red_del.
 DR Pfam; PF02613; Nitrate_red_del; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 191 AA; 22406 MW; 1711F51176028E2F CRC64;

Query Match 21.4%; Score 6; DB 16; Length 191;
 Best Local Similarity 100.0%; Pred. No. 85;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 RAEENI 28
 Db 136 RAEENI 141

RESULT 27

ID Q38495 PRELIMINARY; PRT; 196 AA.
 AC Q38495;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE BACTERIOPHAGE MU LEFT END.
 OS Bacteriophage Mu.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
 OX NCBI_TaxID=10677;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Priess H., Brauer B., Schmidt C., Kamp D.;
 RT "Sequence of the left end of Mu.";
 RL (In) Symonds N., Toussaint A., van de Putte P., Howe M.M. (eds.);
 RL Phage Mu, pp.277-296, Cold Spring Harbor Laboratory, CSH, NY (1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92356834; PubMed=1386645;
 RA Alazard R., Betermier M., Chandler M.;
 RT "Escherichia coli integration host factor stabilizes bacteriophage Mu
 repressor interactions with operator DNA in vitro.";
 RL Mol. Microbiol. 6:1707-1714(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92356835; PubMed=1386646;

RA Gama M.J., Toussaint A., Higgins N.P.;
RT "Stabilization of bacteriophage Mu repressor-operator complexes by the
RT Escherichia coli integration host factor protein.";
RL Mol. Microbiol. 6:1715-1722(1992).
DR EMBL; M64097; AAA32407.1; -.
SQ SEQUENCE 196 AA; 21843 MW; B4B5999C2248AE1F CRC64;

Query Match 21.4%; Score 6; DB 9; Length 196;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 NLP5LC 17
Db 27 NLP5LC 32
|||||

RESULT 28
ID Q9PI81 PRELIMINARY; PRT; 210 AA.
AC Q9PI81;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE PUTATIVE ACIDIC PERIPLASMIC PROTEIN.
GN C30424.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagels K., Kariyasev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrett B.G.;
RT The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
DR EMBL; ALI39075; CAB74260.1; -.
KW Complete proteome.
SQ SEQUENCE 210 AA; 24199 MW; BAD8E66371621107 CRC64;

Query Match 21.4%; Score 6; DB 16; Length 210;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 LRAEEN 27
Db 67 LRAEEN 72
|||||

RESULT 29
ID Q973X2 PRELIMINARY; PRT; 222 AA.
AC Q973X2;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN ST0776.
GN ST0776.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=111955;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,

RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
DR EMBL; AP000983; BAB65788.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 222 AA; 26885 MW; 2D97CD8863AA6FA8 CRC64;

Query Match 21.4%; Score 6; DB 17; Length 222;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ENLPSL 16
Db 131 ENLPSL 136
|||||

RESULT 30
ID Q9C661 PRELIMINARY; PRT; 248 AA.
AC Q9C661;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE HYPOTHETICAL 27.5 KDA PROTEIN.
GN F28B23.16.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Torlumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820(2000).
DR EMBL; AC079829; AAC50684.1; -.
DR InterPro; IPR002819; HD.
DR InterPro; IPR003607; HDC.
DR Pfam; PF01966; HD; 1.
DR SMART; SM00471; HDC; 1.
KW Hypothetical protein.
SQ SEQUENCE 248 AA; 27481 MW; 048FC388AA366442 CRC64;

Query Match 21.4%; Score 6; DB 10; Length 248;
Best Local Similarity 100.0%; Pred. No. 11e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GLRAEE 26

```

Db 165 GLRAEE 170
|||||
RESULT 31
Q97FS0
ID Q97FS0 PRELIMINARY; PRT; 249 AA.
AC Q97FS0;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE POLYFERREDOXIN.
GN CAC2657.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RA "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AF007763; AAK80604.1; -.
DR InterPro; IPR001450; 4Fe4S_ferredoxin.
DR Pfam; PF00037; fer4; 2.
DR PROSITE; PS00198; 4Fe4S_FERREDOXIN; 2.
KW Complete proteome.
SQ SEQUENCE 249 AA; 28892 MW; 1AF664934E46364F CRC64;

Query Match 21.4%; Score 6; DB 16; Length 249;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KENLPS 15
Db 70 KENLPS 75
|||||

RESULT 32
Q93ZV1
ID Q93ZV1 PRELIMINARY; PRT; 258 AA.
AC Q93ZV1;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HYPOTHETICAL 28.7 KDA PROTEIN.
GN F28B23.16.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Banno F., Dale J.M., Goldsmith A.D., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Yamamura Y., Yu G.,
RA Yu S., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C.,
RA Koeseena E., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
RA Southwick A., Tracy S.E., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Full length cDNA of gene F28B23.16 (GI:123221182).";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY056247; AAL07096.1; -.
KW Hypothetical protein.

```

```

SQ SEQUENCE 258 AA; 28695 MW; 5ED5B451E11352A CRC64;

Query Match 21.4%; Score 6; DB 10; Length 258;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GLRAEE 26
Db 175 GLRAEE 180
|||||

RESULT 33
P96452
ID P96452 PRELIMINARY; PRT; 272 AA.
AC P96452;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ORF25 (PUTATIVE SIR2-LIKE TRANSCRIPTIONAL SILENCER PROTEIN).
GN SMB21328.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RM2011;
RX MEDLINE=97175570; PubMed=9023225;
RA Becker A., Rueberg S., Kuester H., Roxlau A.A., Keller M.,
RA Ivashina T., Cheng H., Walker G.C., Puchler A.;
RA "The 32-Kilobase exp gene cluster of Rhizobium meliloti directing the
RT biosynthesis of galactoglucan: genetic organization and properties of
RT the encoded gene products.";
RL J. Bacteriol. 179:1375-1384(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1021; PLASMID=PSYMB (MEGAPLASMID 2);
RX MEDLINE=21396508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA Vorhoelter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA Golding B., Puchler A.;
RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
RT fixing endosymbiont Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL; Z79692; CAB01953.1; -.
DR EMBL; AL603645; CAC49304.1; -.
DR InterPro; IPR003000; SIR2.
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 272 AA; 31324 MW; 287C6B50CF206A25 CRC64;

Query Match 21.4%; Score 6; DB 16; Length 272;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
Db 264 QGLRAE 269
|||||

RESULT 34
Q9KX47
ID Q9KX47 PRELIMINARY; PRT; 287 AA.
AC Q9KX47;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE 2,4-D DIOXYGENASE.
GN TFDA.
OS Alcaligenes denitrificans.
OX Plasmid pEST4011.
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;

```

OC Achromobacter.
OX NCBI_TaxID=32002;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EST4002;
RA Vedler E., Kolv V., Heinaru A.;
RT "Analysis of the 2,4-dichlorophenoxyacetic acid-degradative plasmid
RT pEST4011 of Achromobacter xylosoxidans subsp. denitrificans strain
RT EST4002.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; U321188; AAF59421.1; -.
DR InterPro: IPR003819; Taud_Tfda.
DR Pfam: PF02668; Taud; 1.
KW Dioxxygenase; Plasmid.
SQ SEQUENCE 287 AA; 32585 MW; B8FE70418238DE65 CRC64;

Query Match 21.4%; Score 6; DB 2; Length 287;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
| | | | |
Db 162 QGLRAE 167

RESULT 35
Q93L12 PRELIMINARY; PRT; 287 AA.
AC Q93L12;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE 2,4-D/ALPHA-KETOGLUTARATE DIOXYGENASE.
GN TFDA.
OS Burkholderia cepacia (Pseudomonas cepacia).
OG Plasmid p1JB1.
OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC Burkholderia.
OX NCBI_TaxID=292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2A;
RX MEDLINE=121260388; PubMed=11368091;
RA Poh R.P.C., Xia X., Bruce I.J., Smith A.R.W.;
RT "2,4-dichlorophenoxyacetate/alpha-ketoglutarate dioxxygenases from
RT Burkholderia cepacia 2a and Ralstonia eutropha JMP134.";
RL Microbios 105:43-63(2001).
DR EMBL; AF029344; AAK81681.1; -.
KW Dioxxygenase; Plasmid.
SQ SEQUENCE 287 AA; 32499 MW; B6A6E117522104A6 CRC64;

Query Match 21.4%; Score 6; DB 2; Length 287;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
| | | | |
Db 162 QGLRAE 167

RESULT 36
P96312 PRELIMINARY; PRT; 288 AA.
AC P96312;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE 2,4-D DIOXYGENASE.
GN IJBA.
OS Burkholderia cepacia (Pseudomonas cepacia).
OG Plasmid p1JB.

OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OX Burkholderia.
OX NCBI_TaxID=292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2A;
RA Maclean K.N., Bustin S.A., McClure N.C., Bruce I.J., Smith A.R.W.;
RT "Cloning and Sequence Analysis of the 2,4-D dioxxygenase from plasmid
RT p1JB.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U87394; AAB47567.1; -.
DR InterPro: IPR003819; Taud_Tfda.
DR Pfam: PF02668; Taud; 1.
KW Dioxxygenase; Plasmid.
SQ SEQUENCE 288 AA; 32621 MW; 9D3173629A335C8C CRC64;

Query Match 21.4%; Score 6; DB 2; Length 288;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
| | | | |
Db 163 QGLRAE 168

RESULT 37
Q9FGE1 PRELIMINARY; PRT; 301 AA.
AC Q9FGE1;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE GB|AAD11584.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB025642; BAB08944.1; -.
SQ SEQUENCE 301 AA; 33735 MW; 3EA83D773C6F8A05 CRC64;

Query Match 21.4%; Score 6; DB 10; Length 301;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
| | | | |
Db 117 TGSALT 122

RESULT 38
Q12477 PRELIMINARY; PRT; 304 AA.
AC Q12477;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CHROMOSOME IV READING FRAME ORF YDL179W.
GN PCL9 OR YDL179W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]

RP SEQUENCE FROM N.A.
 RA Pohl T.M.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MIPS;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ALPHAS288C;
 RA Pohl T.M.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY.
 DR EMBL; 274227; CAA98753.1; -.
 DR EMBL; 267750; CAA91566.1; -.
 DR SGD; S0002338; PCL19.
 DR InterPro; IPR004366; Cyclin.
 DR Pfam; PF00134; cyclin; 1.
 DR SMART; SM00385; CYCLIN; 1.
 KW Cell cycle; Cell division; Cyclin; Hypothetical protein.
 SQ SEQUENCE 304 AA; 34798 MW; 183820067EC00D23 CRC64;

Query Match 21.4%; Score 6; DB 3; Length 304;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 RAEENI 28
 |||||
 Db 260 RAEENI 265

RESULT 39
 O30246
 ID O30246 PRELIMINARY; PRT; 324 AA.
 AC O30246;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE METHANOL DEHYDROGENASE REGULATORY PROTEIN (MOXR).
 GN AF2425.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
 OC Archaeoglobus.
 OC NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Winn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyripides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spraggins T., Artiach P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 reducing archaeon Archaeoglobus fulgidus."
 RL Nature 390:364-370(1997).
 DR EMBL; AE001108; AAB91247.1; -.
 DR TIGR; AF2425; -.
 DR InterPro; IPR000523; Mg_chelatase_chiI.
 DR Pfam; PF01078; Mg_chelatase; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 324 AA; 36609 MW; B0F64965EFC82F1F CRC64;

Query Match 21.4%; Score 6; DB 17; Length 324;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 21 GLRAEE 26
 |||||
 Db 300 GLRAEE 305

RESULT 40
 Q9NWE9
 ID Q9NWE9 PRELIMINARY; PRT; 332 AA.
 AC Q9NWE9;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE HYPOTHETICAL 37.8 KDA PROTEIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=EMBRYO;
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
 RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
 RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
 RA Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
 RT "NEDO human cDNA sequencing project";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK000936; BAA91433.1; -.
 DR InterPro; IPR000719; Euk_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Transferase.
 SQ SEQUENCE 332 AA; 37848 MW; A9E7734E0C1FB86E CRC64;

Query Match 21.4%; Score 6; DB 4; Length 332;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 NLPSLC 17
 |||||
 Db 48 NLPSLC 53

RESULT 41
 Q9RDZ8
 ID Q9RDZ8 PRELIMINARY; PRT; 339 AA.
 AC Q9RDZ8;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PUTATIVE GLYCOSYL TRANSFERASE.
 OS Legionella pneumophila.
 OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
 OC Legionellaceae; Legionella.
 OC NCBI_TaxID=446;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=RCL;
 RX MEDLINE=200496696; PubMed=11043980;
 RA Lueneberg E., Zetzmann N., Hartmann M., Knirel Y.A., Koolstra O.,
 RA Zaehring U., Helbig J., Frosch M.;
 RT "Cloning and functional characterization of a 30 kb gene locus
 required for lipopolysaccharide biosynthesis in Legionella
 pneumophila."
 RL Int. J. Med. Microbiol. 290:37-49(2000).
 DR EMBL; AJ007311; CAB65189.1; -.
 DR InterPro; IPR001173; Glycos_transf_2.
 DR Pfam; PF00535; Glycos_transf_2; 1.
 KW Transferase
 SQ SEQUENCE 339 AA; 38892 MW; 71DE40EB34F7DE9E CRC64;

```

Query Match      21.4%; Score 6; DB 2; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
Db 237 TGSALT 242

RESULT 42
Q966M2
ID Q966M2 PRELIMINARY; PRT; 346 AA.
AC Q966M2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN C55B7.3.
GN C55B7.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Wansley P.;
RT "The sequence of C. elegans cosmid C55B7.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006625; AAK68274.1; -.
SQ SEQUENCE 346 AA; 39847 MW; B0795B6C24276D2B CRC64;

Query Match      21.4%; Score 6; DB 5; Length 346;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
Db 54 QGLRAE 59

RESULT 43
Q9XVU4
ID Q9XVU4 PRELIMINARY; PRT; 352 AA.
AC Q9XVU4;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE F10G8.1 PROTEIN.
GN F10G8.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Basham V.M.;

Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: A platform for
investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z80216; CAB02287.1; -.
DR HSSP; P18052; LYFO.
DR InterPro; IPR000387; TYR_phosphatase.
DR InterPro; IPR000242; TYR_prot_phptase.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PRO0700; PRTYPHPHTASE.
DR SMART; SM00194; PTPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE.1; 1.
DR PROSITE; PS00056; TYR_PHOSPHATASE.2; 1.
DR PROSITE; PS00055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase.
SQ SEQUENCE 352 AA; 40502 MW; C448B4442412C70C CRC64;

Query Match      21.4%; Score 6; DB 5; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
Db 54 QGLRAE 59

RESULT 44
Q90634
ID Q90634 PRELIMINARY; PRT; 359 AA.
AC Q90634;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE TOP AP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-POSTERIOR QUADRANT RETINA;
RX MEDLINE=95161057; PubMed=7857637;
RA Savitt J.M., Trisler D., Hilt D.C.;
RT "Molecular cloning of TOPAP: a topographically graded protein in the
developing chick visual system.";
RL Neuron 14:253-261(1995).
DR EMBL; U17000; AAA68637.1; -.
SQ SEQUENCE 359 AA; 41425 MW; A580146CF28DC01D CRC64;

Query Match      21.4%; Score 6; DB 13; Length 359;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 QCKENL 13
Db 318 QCKENL 323

RESULT 45
Q98DU5
ID Q98DU5 PRELIMINARY; PRT; 366 AA.
AC Q98DU5;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE OUTER MEMBRANE LIPOPROTEIN GNA33, MEMBRANE-BOUND LYTIC MUREIN
```

DE TRANSLYCOSYLASE.
GN MLR4546.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Rhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL: AP003004; BAB51175.1; -;
KW Lipoprotein; Complete proteome.
SQ SEQUENCE 366 AA; 39892 MW; 75579DAA73E8E4EC CRC64;

Query Match 21.4%; Score 6; DB 16; Length 366;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSAT 6
|||||
Db 320 TGSAT 325

RESULT 46
Q96Q98 PRELIMINARY; PRT; 377 AA.
AC Q96Q98;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DE FKSG10.
GN FKSG10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang Y.-G.;
RT "Identification of FKSG10, a novel gene related to breast cancer.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY008304; AAG22861.1; -;
SQ SEQUENCE 377 AA; 41534 MW; 8CEDAD564E8680AD CRC64;

Query Match 21.4%; Score 6; DB 4; Length 377;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
|||||
Db 288 QGLRAE 293

RESULT 47
Q9N2H6 PRELIMINARY; PRT; 386 AA.
AC Q9N2H6;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DE SELENOPROTEIN P.
GN SELP.
OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Hara S., Imura N., Shoji Y.;
RT "Bovine endothelial selenoprotein P.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB032826; BAA84781.1; -;
SQ SEQUENCE 386 AA; 43061 MW; EC6F3A76874E528F CRC64;

Query Match 21.4%; Score 6; DB 6; Length 386;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSAT 6
|||||
Db 316 TGSAT 321

RESULT 48
Q49722 PRELIMINARY; PRT; 394 AA.
AC Q49722;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE HYPOTHETICAL 40.8 KDA PROTEIN ML0405.
GN ML0405 OR B1620_C2_213 OR MLCL383.01.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Smith D.R., Robison K.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 231-394 FROM N.A.
RA Oliver K., Harris D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 231-394 FROM N.A.
RX MEDLINE=93188700; PubMed=8446027;
RA Eiglmeier K., Honore N., Woods S.A., Caudron B., Cole S.T.;
RT "Use of an ordered cosmid library to deduce the genomic organization
of Mycobacterium leprae.";
RL Mol. Microbiol. 7:197-206(1993).
DR EMBL: AL583918; CAC29913.1; -;
DR EMBL: U00015; AAC43223.1; -;
DR EMBL: Z97179; CAB09940.1; -;
DR Leproma; ML0405; -;
DR Hypothetical protein; Transmembrane; Complete proteome.
KW TRANSMEM 279 299 POTENTIAL.
SQ SEQUENCE 394 AA; 40754 MW; DOE455C74ED7A10C CRC64;

Query Match 21.4%; Score 6; DB 16; Length 394;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ENLPSL 16
| | | | |
DB 274 ENLPSL 279

RESULT 49

Q987Z2 PRELIMINARY; PRT; 411 AA.
AC Q987Z2;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MLL6844 PROTEIN.
GN MLL6844.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti".
RL DNA Res 7:331-338(2000).
DR EMBL; AF003010; BAB53058.1; -.
KW Complete proteome.
SQ SEQUENCE 411 AA; 43681 MW; 506DFBE4247ABCA3 CRC64;

Query Match 21.4%; Score 6; DB 16; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
| | | | |
DB 349 TGSALT 354

RESULT 50

P78356 PRELIMINARY; PRT; 416 AA.
AC P78356;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PHOSPHATIDYLINOSITOL-4-PHOSPHATE 5-KINASE TYPE II BETA (EC
2.7.1.68).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Castellino A.M., Parker G.J., Boronkov I.V., Anderson R.A.,
RA Chao M.V.;
RT "A Novel Interaction between the Juxtamembrane Region of the p55 Tumor
Necrosis Factor Receptor and Phosphatidylinositol-4-phosphate 5-
Kinase".
RT Kinase".
RL J. Biol. Chem. 272:0-0(1997).
DR EMBL; U85245; AAB48596.1; -.
DR InterPro; IPR002498; PIP5K.

DR Pfam; PF01504; PIP5K; 1.
DR SMART; SM00330; PIP5K; 1.
KW Kinase; Transferase.
SQ SEQUENCE 416 AA; 47378 MW; 2288CD2883EACEBE CRC64;

Query Match 21.4%; Score 6; DB 4; Length 416;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KENLPS 15
| | | | |
DB 86 KENLPS 91

Search completed: August 22, 2002, 11:20:57
Job time: 244 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 11:19:15 ; Search time 52.19 Seconds
(without alignments)
59.591 Million cell updates/sec

Title: US-09-856-199-2

Perfect score: 28

Sequence: 1 TGSATYQCKENLPSCSXOGLRAEENI 28

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1200 summaries

Database : A_Geneseq_032802.*

1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	11	39.3	20	21	AA03189 Human selenoprotei
2	11	39.3	28	21	AB03184 Human selenoprotei
3	11	39.3	103	21	AA03185 Human selenoprotei
4	11	39.3	381	21	AA03188 Human selenoprotei
5	7	25.0	305	22	AA075466 Human colon cancer
6	7	25.0	832	22	AB023820 Novel human diagno
7	7	25.0	932	22	AB017156 Novel human diagno
8	6	21.4	62	22	AA064969 Propionibacterium
9	6	21.4	62	22	AA06318 Human foetal prote
10	6	21.4	97	22	AA067435 Propionibacterium
11	6	21.4	107	22	AA031563 Novel human secret

12	6	21.4	120	22	AAU52347 Propionibacterium
13	6	21.4	166	22	AB062732 Drosophila melanog
14	6	21.4	197	21	AA06557 Arabidopsis thalia
15	6	21.4	238	21	AA06557 Arabidopsis thalia
16	6	21.4	266	22	AA06556 Wheat cis-prenyltr
17	6	21.4	273	21	AA06555 Arabidopsis thalia
18	6	21.4	280	17	AAW01142 TrpA protein. Bre
19	6	21.4	280	22	AAW01142 Human polypeptide
20	6	21.4	280	22	AAW01142 C glutamicum prote
21	6	21.4	280	22	AAW01142 Corynebacterium gl
22	6	21.4	281	8	AAW01142 TrpA enzyme. AAP
23	6	21.4	310	20	AAW01142 Human prostate tum
24	6	21.4	332	22	AAW01142 Human protein sequ
25	6	21.4	399	21	AAW01142 Arabidopsis thalia
26	6	21.4	404	21	AAW01142 Human secreted pro
27	6	21.4	431	21	AAW01142 Arabidopsis thalia
28	6	21.4	442	10	AAW01142 N-terminal dimer o
29	6	21.4	443	22	AAW01142 Human peptidyl-pro
30	6	21.4	469	22	AAW01142 Drosophila melanog
31	6	21.4	480	21	AAW01142 Synchocystis sp.
32	6	21.4	481	21	AAW01142 Synchocystis sp p
33	6	21.4	482	22	AAW01142 Drosophila melanog
34	6	21.4	484	22	AAW01142 Human protein sequ
35	6	21.4	489	21	AAW01142 Human secreted pro
36	6	21.4	492	22	AAW01142 Protein encoded by
37	6	21.4	505	21	AAW01142 Novel protein kina
38	6	21.4	508	19	AAW01142 Fibroblast growth
39	6	21.4	508	21	AAW01142 Escherichia coli y
40	6	21.4	508	22	AAW01142 Human suc-1-associ
41	6	21.4	516	21	AAW01142 Human GTPase assoc
42	6	21.4	521	22	AAW01142 Human FGFR signal
43	6	21.4	544	21	AAW01142 Human PRO1374 (UNQ
44	6	21.4	544	22	AAW01142 Human PRO polypept
45	6	21.4	544	22	AAW01142 Protein of the inv
46	6	21.4	544	22	AAW01142 Amino acid sequenc
47	6	21.4	549	22	AAW01142 Human novel secret
48	6	21.4	551	16	AAW01142 Cattle MIS mature
49	6	21.4	557	21	AAW01142 Arabidopsis thalia
50	6	21.4	575	8	AAW01142 Arabidopsis thalia
51	6	21.4	575	10	AAW01142 Bovine Mullerian i
52	6	21.4	575	16	AAW01142 Cattle MIS protein
53	6	21.4	632	22	AAW01142 Novel human diagno
54	6	21.4	709	19	AAW01142 Homo sapiens fetal
55	6	21.4	709	22	AAW01142 Human BD335_14 pro
56	6	21.4	735	22	AAW01142 Human protein sequ
57	6	21.4	756	22	AAW01142 Human protein sequ
58	6	21.4	756	22	AAW01142 Amino acid sequenc
59	6	21.4	759	22	AAW01142 Human protein sequ
60	6	21.4	792	22	AAW01142 Human kinase PKIN-
61	6	21.4	847	22	AAW01142 Novel human diagno
62	6	21.4	847	22	AAW01142 Novel human diagno
63	6	21.4	851	22	AAW01142 Novel human diagno
64	6	21.4	873	22	AAW01142 Drosophila melanog
65	6	21.4	1083	22	AAW01142 Human potassium-ch
66	6	21.4	1083	22	AAW01142 Mouse potassium-ch
67	6	21.4	1083	22	AAW01142 Mouse mKCC4 mutant
68	6	21.4	1083	22	AAW01142 Mouse mKCC4 mutant
69	6	21.4	1083	22	AAW01142 Human protein SEQ
70	6	21.4	1088	21	AAW01142 Human ORFX ORF2170
71	6	21.4	1090	22	AAW01142 Human K/Cl cotrans
72	6	21.4	1090	22	AAW01142 Human protein SEQ
73	6	21.4	1134	22	AAW01142 Novel human diagno
74	6	21.4	1297	22	AAW01142 Drosophila melanog
75	6	21.4	1420	18	AAW01142 Yeast transcriptio
76	6	21.4	1507	22	AAW01142 Drosophila melanog
77	6	21.4	1621	22	AAW01142 Human protein SEQ
78	6	21.4	1644	22	AAW01142 Novel human diagno
79	6	21.4	2293	21	AAW01142 Amino acid sequenc
80	6	21.4	3868	22	AAW01142 Drosophila melanog
81	5	17.9	10	22	AAW01142 Human complementar
82	5	17.9	13	16	AAW01142 Human MHC class II
83	5	17.9	16	20	AAW01142 Rheumatoid arthrit
84	5	17.9	16	20	AAW01142 Rheumatoid arthrit

85	5	17.9	16	22	ABB52233	Human API-181 trypt	158	5	17.9	67	22	AAU22683	Novel human colon
86	5	17.9	16	22	ABB52288	Human API-80 trypt	159	5	17.9	67	22	AAU93129	Human digestive sy
87	5	17.9	16	22	ABB52314	Human API-42 trypt	160	5	17.9	68	18	AAU20460	H. pylori secreted
88	5	17.9	16	22	ABB52393	Human API-140 trypt	161	5	17.9	68	22	ABB15238	Human nervous syst
89	5	17.9	16	22	AAU25030	Schizophrenia-Asso	162	5	17.9	69	20	AAU12170	Human 5' EST secre
90	5	17.9	16	22	AAU15374	Schizophrenia-Asso	163	5	17.9	71	13	AAU20985	Human cancer assoc
91	5	17.9	17	22	AAU08978	Human CDC25A inhib	164	5	17.9	72	21	AAU44166	Non-A, non-B hepat
92	5	17.9	19	22	ABB31067	Peptide #3718 enco	165	5	17.9	72	21	AAU35754	Human cancer assoc
93	5	17.9	19	22	ABB36256	Peptide #3762 enco	166	5	17.9	73	22	AAU02131	Zea mays protein f
94	5	17.9	19	22	AAU57031	Human brain expres	167	5	17.9	73	22	AAU45563	Novel human diagno
95	5	17.9	19	22	AAU72403	Human bone marrow	168	5	17.9	74	22	AAU45563	Propionibacterium
96	5	17.9	19	22	AAU17257	Peptide #3691 enco	169	5	17.9	75	22	AAU45563	Propionibacterium
97	5	17.9	19	22	AAU29752	Peptide #3789 enco	170	5	17.9	75	22	AAU29752	Propionibacterium
98	5	17.9	22	22	ABG07762	Novel human diagno	171	5	17.9	77	20	AAU21489	Human normal bladd
99	5	17.9	26	20	AAU88752	Secreted protein e	172	5	17.9	77	22	AAU37442	Staphylococcus aur
100	5	17.9	26	22	ABU50527	Human secreted pro	173	5	17.9	77	22	ABU50527	Novel human diagno
101	5	17.9	30	16	AAU74057	Superantigen pepti	174	5	17.9	79	21	AAU54421	Human pancreatic c
102	5	17.9	31	13	AAU39977	Hepatitis C virus	175	5	17.9	80	22	AAU58388	Propionibacterium
103	5	17.9	31	22	AAU2286	Hepatitis C virus	176	5	17.9	81	21	AAU24850	Arabidopsis thalia
104	5	17.9	31	22	AAU37510	Peptide #5016 enco	177	5	17.9	81	21	AAU26162	Zea mays protein f
105	5	17.9	31	22	AAU58172	Human brain expres	178	5	17.9	81	21	AAU30837	Arabidopsis thalia
106	5	17.9	31	22	AAU18469	Peptide #4903 enco	179	5	17.9	81	22	AAU57229	Propionibacterium
107	5	17.9	34	22	ABU39271	Peptide #6777 enco	180	5	17.9	83	22	AAU51473	Propionibacterium
108	5	17.9	34	22	AAU59335	Human brain expres	181	5	17.9	84	22	ABG29267	Novel human diagno
109	5	17.9	34	22	AAU72530	Human bone marrow	182	5	17.9	84	22	AAU00848	Human bone marrow
110	5	17.9	35	22	AAU3677	Peptide #2359 enco	183	5	17.9	85	22	AAU00848	Novel human diagno
111	5	17.9	35	22	AAU32774	Peptide #6811 enco	184	5	17.9	86	21	AAU55571	Arabidopsis thalia
112	5	17.9	35	22	ABU34933	Peptide #2439 enco	185	5	17.9	86	22	ABU15670	Human nervous syst
113	5	17.9	35	22	ABU20344	Protein #2343 enco	186	5	17.9	87	20	AAU14440	Human secreted pro
114	5	17.9	35	22	AAU22021	Human cardiovascular	187	5	17.9	88	19	AAU77609	Staphylococcus aur
115	5	17.9	35	22	AAU68120	Human bone marrow	188	5	17.9	88	22	AAU90638	Human immune/haema
116	5	17.9	35	22	AAU3677	Peptide #2359 enco	189	5	17.9	88	22	AAU06289	Human foetal prote
117	5	17.9	38	11	AAU07516	Portion of Trichos	190	5	17.9	88	22	AAU06704	Human foetal prote
118	5	17.9	40	21	AAU45865	Arabidopsis thalia	191	5	17.9	89	12	AAU12159	NANB hepatitis vir
119	5	17.9	40	22	ABU42432	Peptide #9938 enco	192	5	17.9	89	22	AAU44512	Propionibacterium
120	5	17.9	40	22	ABU25882	Protein #7881 enco	193	5	17.9	89	22	AAU67315	Propionibacterium
121	5	17.9	40	22	AAU63318	Human brain expres	194	5	17.9	92	21	AAU08597	Arabidopsis thalia
122	5	17.9	40	22	AAU76133	Human bone marrow	195	5	17.9	92	21	AAU48272	Arabidopsis thalia
123	5	17.9	42	22	AAU36243	Peptide #10280 enc	196	5	17.9	92	22	AAU39447	Propionibacterium
124	5	17.9	42	21	AAU46071	Arabidopsis thalia	197	5	17.9	92	22	AAU48446	Propionibacterium
125	5	17.9	44	20	AAU00634	Human telomerase p	198	5	17.9	92	22	ABU16796	Novel human diagno
126	5	17.9	45	22	AAU82809	Human immune/haema	199	5	17.9	93	13	AAU20566	Non-A non-B hepati
127	5	17.9	45	22	AAU73518	Human colon cancer	200	5	17.9	93	21	AAU51630	Human secreted pro
128	5	17.9	47	22	ABU29739	Peptide #2390 enco	201	5	17.9	93	22	AAU32720	Novel human secret
129	5	17.9	47	22	ABU34910	Peptide #2416 enco	202	5	17.9	95	21	AAU08596	Arabidopsis thalia
130	5	17.9	47	22	ABU20324	Protein #2323 enco	203	5	17.9	95	21	AAU48271	Arabidopsis thalia
131	5	17.9	47	22	AAU55721	Human brain expres	204	5	17.9	96	13	AAU25436	Tapetum protein fr
132	5	17.9	47	22	AAU68099	Human bone marrow	205	5	17.9	96	22	AAU89501	Human immune/haema
133	5	17.9	47	22	AAU15923	Peptide #2357 enco	206	5	17.9	97	22	AAU27509	Human G-protein Co
134	5	17.9	47	22	AAU8426	Peptide #2463 enco	207	5	17.9	97	22	AAU07822	Human polypeptide
135	5	17.9	47	22	AAU03659	Peptide #2341 enco	208	5	17.9	98	22	AAU42381	Propionibacterium
136	5	17.9	48	22	AAU03554	Human trefol doma	209	5	17.9	100	21	AAU37854	Neisserial consery
137	5	17.9	49	21	AAU65212	Human 5' EST relat	210	5	17.9	100	21	AAU07849	Arabidopsis thalia
138	5	17.9	50	22	AAU65180	Propionibacterium	211	5	17.9	102	15	AAU48563	Sequence of the co
139	5	17.9	52	22	AAU53961	Propionibacterium	212	5	17.9	102	22	AAU10686	Human polypeptide
140	5	17.9	52	22	ABU03907	Human musculoskele	213	5	17.9	103	13	AAU29872	HCV NS4-NS5 peptid
141	5	17.9	53	22	AAU76593	Human colon cancer	214	5	17.9	103	13	AAU29887	HCV NS4-NS5 peptid
142	5	17.9	54	22	ABU6366	Novel human diagno	215	5	17.9	103	13	AAU29888	HCV NS4-NS5 peptid
143	5	17.9	55	22	AAU48253	Propionibacterium	216	5	17.9	103	13	AAU29889	HCV NS4-NS5 peptid
144	5	17.9	56	21	AAU54241	Human pancreatic c	217	5	17.9	103	13	AAU29890	HCV NS4-NS5 peptid
145	5	17.9	57	22	AAU85515	Human protein kina	218	5	17.9	103	13	AAU29891	Propionibacterium
146	5	17.9	58	22	AAU12853	Human 5' EST secre	219	5	17.9	103	22	AAU40887	Propionibacterium
147	5	17.9	58	21	AAU16808	Bacteriophage Dp-1	220	5	17.9	103	22	AAU62993	Human immune/haema
148	5	17.9	59	22	AAU45611	Propionibacterium	221	5	17.9	103	22	AAU90462	Human immune/haema
149	5	17.9	59	22	AAU59104	Propionibacterium	222	5	17.9	103	22	AAU93363	Human polypeptide,
150	5	17.9	59	22	AAU64273	Propionibacterium	223	5	17.9	103	22	AAU23872	M pneumoniae Est e
151	5	17.9	61	21	AAU44209	Human cancer assoc	224	5	17.9	105	22	AAU57708	Propionibacterium
152	5	17.9	61	22	AAU48746	Propionibacterium	225	5	17.9	105	22	AAU12248	Human PRO4381 poly
153	5	17.9	61	22	AAU86110	Human immune/haema	226	5	17.9	106	21	AAU37047	Arabidopsis thalia
154	5	17.9	61	22	AAU03856	Human polypeptide	227	5	17.9	106	21	AAU47491	Arabidopsis thalia
155	5	17.9	61	22	AAU77203	Human colon cancer	228	5	17.9	106	21	AAU47520	Arabidopsis thalia
156	5	17.9	62	22	AAU85432	Human immune/haema	229	5	17.9	106	22	AAU60002	Propionibacterium
157	5	17.9	65	22	AAU64925	Propionibacterium	230	5	17.9	107	22	AAU82547	Human immune/haema
												AAU04050	Plant defensin. p

231	5	17.9	107	22	AAB04051	Plant defensin. P	304	5	17.9	137	21	AAB08207	CDNA encoding a hu
232	5	17.9	107	22	AAB04053	Plant defensin. P	305	5	17.9	139	21	AAU49605	Propionibacterium
233	5	17.9	107	22	AAB04054	Plant defensin. P	306	5	17.9	140	21	AAG08457	Arabidopsis thalia
234	5	17.9	107	22	AAB04056	Plant defensin. P	307	5	17.9	140	21	AAG23471	Arabidopsis thalia
235	5	17.9	108	22	AAU45900	Propionibacterium	308	5	17.9	140	21	AAG24848	Arabidopsis thalia
236	5	17.9	108	22	AAB61780	Sunflower defensin	309	5	17.9	140	21	AAG26160	Zea mays protein f
237	5	17.9	108	22	AAB61795	Sunflower defensin	310	5	17.9	140	21	AAG03817	Human secreted pro
238	5	17.9	109	21	AAG48270	Arabidopsis thalia	311	5	17.9	140	22	ABB60072	Drosophila melanog
239	5	17.9	109	21	AAG55570	Arabidopsis thalia	312	5	17.9	140	22	ABB10165	Human cDNA SEQ ID
240	5	17.9	109	22	AAO06010	Human polypeptide	313	5	17.9	140	22	AAU18096	Novel human uterin
241	5	17.9	109	22	AAG77352	Human colon cancer	314	5	17.9	140	22	AAW78591	Human protein SEQ
242	5	17.9	110	21	AAB36607	Human prostate can	315	5	17.9	140	22	AAU16995	Human novel secret
243	5	17.9	110	21	AAB08595	Arabidopsis thalia	316	5	17.9	140	22	AAU19908	Novel human calciu
244	5	17.9	111	21	AAG76165	Zea mays protein f	317	5	17.9	141	22	ABB67536	Drosophila melanog
245	5	17.9	111	22	ABB27728	Human peptide #379	318	5	17.9	142	22	AAO09303	Human polypeptide
246	5	17.9	111	22	ABB32899	Peptide #405 encod	319	5	17.9	143	22	AAU27410	Novel bone marrow
247	5	17.9	111	22	ABB18379	Protein #378 encod	320	5	17.9	143	22	AAO09501	Human polypeptide
248	5	17.9	111	22	AAW53701	Human brain expres	321	5	17.9	144	21	AAG17576	Arabidopsis thalia
249	5	17.9	111	22	AAW66084	Human bone marrow	322	5	17.9	144	21	AAG23470	Arabidopsis thalia
250	5	17.9	111	22	AAW13954	Peptide #388 encod	323	5	17.9	144	21	AAG43066	Arabidopsis thalia
251	5	17.9	111	22	AAW26360	Peptide #397 encod	324	5	17.9	144	21	AAG59705	Arabidopsis thalia
252	5	17.9	111	22	AAW01697	Peptide #379 encod	325	5	17.9	145	21	AAG26163	Zea mays protein f
253	5	17.9	112	20	AAW48247	Human prostate can	326	5	17.9	145	22	ABG20115	Novel human diagno
254	5	17.9	112	22	AAW32855	Novel human secret	327	5	17.9	146	13	AAW23882	NANBH peptide A.
255	5	17.9	113	22	AAW85250	Human immune/haema	328	5	17.9	146	21	AAG16949	Arabidopsis thalia
256	5	17.9	113	22	AAU17303	Novel signal trans	329	5	17.9	146	21	AAG01595	Human secreted pro
257	5	17.9	113	22	AAU17614	Novel signal trans	330	5	17.9	147	22	ABG10170	Novel human diagno
258	5	17.9	113	22	AAU02514	Anti-adipocyte mon	331	5	17.9	148	13	AAW26410	Human phospholipas
259	5	17.9	114	22	AAG89914	C glutamicum prote	332	5	17.9	148	16	AAW63054	Human PLA2 type I.
260	5	17.9	114	22	AAE03550	Human trefol doma	333	5	17.9	149	21	AAW59704	Arabidopsis thalia
261	5	17.9	115	20	AAW39548	Human CTLA-4 VDL s	334	5	17.9	151	21	AAW59845	Arabidopsis thalia
262	5	17.9	115	22	AAB04055	Plant defensin. P	335	5	17.9	151	21	AAG60285	Arabidopsis thalia
263	5	17.9	116	20	AAW12323	Human 5' EST secre	336	5	17.9	151	22	AAU52565	Propionibacterium
264	5	17.9	116	21	AAW16951	Arabidopsis thalia	337	5	17.9	152	22	AAU63526	Propionibacterium
265	5	17.9	117	20	AAW35467	Chlamydia pneumoni	338	5	17.9	152	22	ABG27571	Novel human diagno
266	5	17.9	117	21	AAW07580	Protein encoded by	339	5	17.9	152	22	AAW83746	Human immune/haema
267	5	17.9	117	22	AAU44579	Propionibacterium	340	5	17.9	153	22	ABG06304	Novel human diagno
268	5	17.9	118	21	AAW08460	Rat Ptx2 protein.	341	5	17.9	154	22	AAW79575	Human protein SEQ
269	5	17.9	118	21	AAG22937	Arabidopsis thalia	342	5	17.9	154	22	AAO03682	Human polypeptide
270	5	17.9	119	21	AAW40577	Human ORFX ORF341	343	5	17.9	155	22	AAU30076	Novel human secret
271	5	17.9	119	22	AAU02526	Anti-adipocyte mon	344	5	17.9	156	21	AAW54059	Human pancreatic c
272	5	17.9	120	21	AAW93700	Cytotoxic T-lympho	345	5	17.9	156	21	AAG37046	Arabidopsis thalia
273	5	17.9	120	22	AAW62156	Propionibacterium	346	5	17.9	156	21	AAG47490	Arabidopsis thalia
274	5	17.9	120	22	AAW83436	Human immune/haema	347	5	17.9	156	21	AAG47519	Arabidopsis thalia
275	5	17.9	121	20	AAW40837	SCA a recombinant	348	5	17.9	156	22	ABW27680	Human peptide #331
276	5	17.9	121	21	AAW29930	Human CTLA-4 scaff	349	5	17.9	156	22	ABW27680	Human peptide #331
277	5	17.9	121	22	ABG26508	Novel human diagno	350	5	17.9	156	22	ABW18332	Peptide #356 encod
278	5	17.9	121	22	ABW16886	Human nervous syst	351	5	17.9	156	22	AAW66036	Human bone marrow
279	5	17.9	121	22	AAU16326	Human novel secret	352	5	17.9	156	22	AAW13905	Peptide #339 encod
280	5	17.9	121	18	AAW35847	Human CTLA4 for us	353	5	17.9	156	22	AAW26311	Peptide #348 encod
281	5	17.9	124	21	AAW16950	Arabidopsis thalia	354	5	17.9	156	22	AAW01648	Peptide #330 encod
282	5	17.9	125	19	AAW52232	Antibody LD2-14-VH	355	5	17.9	157	21	AAG30835	Arabidopsis thalia
283	5	17.9	125	19	AAW52240	Antibody LD1/2-6-3	356	5	17.9	157	21	AAG54841	Arabidopsis thalia
284	5	17.9	125	19	AAW52242	Antibody LD1/2-6-3	357	5	17.9	159	12	AAW10759	Non-A non-B hepati
285	5	17.9	125	21	AAG08458	Arabidopsis thalia	358	5	17.9	159	22	ABW27842	Novel human diagno
286	5	17.9	125	21	AAG23472	Arabidopsis thalia	359	5	17.9	160	20	AAW06399	Mouse NTR-5 recept
287	5	17.9	125	21	AAG24849	Arabidopsis thalia	360	5	17.9	160	22	AAW95531	Human protein sequ
288	5	17.9	125	21	AAG26161	Zea mays protein f	361	5	17.9	161	20	AAW60048	Human endometrium
289	5	17.9	125	21	AAG30836	Arabidopsis thalia	362	5	17.9	164	22	ABW18934	Novel human diagno
290	5	17.9	126	21	AAG26164	Zea mays protein f	363	5	17.9	164	22	AAW78333	Human zlecl protei
291	5	17.9	127	18	AAW28273	Staphylococcus aur	364	5	17.9	165	21	AAG08456	Arabidopsis thalia
292	5	17.9	127	20	AAW66152	Human bladder tumo	365	5	17.9	165	22	ABW52688	Escherichia coli p
293	5	17.9	127	22	ABG32099	Novel human diagno	366	5	17.9	168	19	AAW42340	CTLA-4 extracellul
294	5	17.9	128	22	AAW79697	Human protein SEQ	367	5	17.9	168	21	AAG25103	Arabidopsis thalia
295	5	17.9	129	22	AAW76592	Corynebacterium gl	368	5	17.9	168	21	AAG26366	Arabidopsis thalia
296	5	17.9	131	22	AAO02207	Human polypeptide	369	5	17.9	170	21	AAG25106	Arabidopsis thalia
297	5	17.9	132	22	AAW78323	Human protein SEQ	370	5	17.9	170	21	AAG26365	Arabidopsis thalia
298	5	17.9	133	22	AAW32290	Novel human secret	371	5	17.9	172	22	AAU52116	Propionibacterium
299	5	17.9	133	22	AAE03253	Human gene 4 encod	372	5	17.9	173	21	AAG22434	Arabidopsis thalia
300	5	17.9	134	22	AAU40131	Propionibacterium	373	5	17.9	174	12	AAW10752	Non-A non-B hepati
301	5	17.9	135	22	ABW56764	Drosophila melanog	374	5	17.9	174	21	AAW08208	Amino acid sequenc
302	5	17.9	136	21	AAB43283	Human ORFX ORF3047	375	5	17.9	174	22	AAW95201	Human reproductive
303	5	17.9	136	22	AAW48080	Human extracellula	376	5	17.9	175	12	AAW10746	Non-A non-B hepati

377	5	17.9	175	13	AAR20565	Non-A non-B hepati	450	5	17.9	212	21	AAG34465	Arabidopsis thalia
378	5	17.9	175	21	AAG17575	Arabidopsis thalia	451	5	17.9	212	21	AAG52142	Arabidopsis thalia
379	5	17.9	175	21	AAG43065	Arabidopsis thalia	452	5	17.9	213	22	AAU29495	Human G-protein-co
380	5	17.9	175	22	ABG14706	Novel human diagno	453	5	17.9	214	20	AAU06522	Mouse STRIFE1 (Tan
381	5	17.9	176	21	AAG20142	Arabidopsis thalia	454	5	17.9	214	20	AAW98145	Mouse TRAIN-R (lon
382	5	17.9	176	21	AAG30951	Arabidopsis thalia	455	5	17.9	214	20	AAW93580	Mouse MAP04-alpha
383	5	17.9	176	21	AAG48902	Arabidopsis thalia	456	5	17.9	214	21	AAW23548	Murine dtroy prote
384	5	17.9	176	21	AAG48910	Arabidopsis thalia	457	5	17.9	214	21	AAW61446	Arabidopsis thalia
385	5	17.9	178	22	AAU34370	Staphylococcus aur	458	5	17.9	214	21	AAW91493	Human secreted pro
386	5	17.9	178	22	AAU36627	Staphylococcus aur	459	5	17.9	214	22	AAW63854	Human prostate can
387	5	17.9	178	22	AAU37237	Staphylococcus aur	460	5	17.9	215	21	AAW43741	Arabidopsis thalia
388	5	17.9	179	20	AAU07009	Breast cancer asso	461	5	17.9	215	22	ABG00152	Novel human diagno
389	5	17.9	179	21	AAU09630	Arabidopsis thalia	462	5	17.9	216	22	ABG72831	Human olfactory re
390	5	17.9	179	21	AAW61435	Arabidopsis thalia	463	5	17.9	217	22	ABG11912	Novel human diagno
391	5	17.9	182	19	AAW60985	Streptococcus pneu	464	5	17.9	219	13	AAW20818	CD53 haematopoieti
392	5	17.9	183	22	ABW67484	Drosophila melanog	465	5	17.9	219	17	AAW91446	Human CD53 antigen
393	5	17.9	184	12	AAW10750	Non-A non-B hepati	466	5	17.9	219	19	AAW80455	Human CD53 antigen
394	5	17.9	184	21	AAW21697	Arabidopsis thalia	467	5	17.9	219	20	AAW89152	Human CD53 antigen
395	5	17.9	184	22	AAU27483	Human G-protein Co	468	5	17.9	219	21	AAW96141	Human haematopoiet
396	5	17.9	186	22	ABG27186	Novel human diagno	469	5	17.9	219	22	AAU02450	Human lymphocyte c
397	5	17.9	186	22	ABW53060	Escherichia coli p	470	5	17.9	220	21	AAW07848	Amino acid sequenc
398	5	17.9	186	22	AAW74651	Human colon cancer	471	5	17.9	220	22	ABG14815	Novel human diagno
399	5	17.9	187	17	AAW86945	Human T cell speci	472	5	17.9	221	21	AAW07850	Novel acid sequenc
400	5	17.9	187	19	AAW29728	Soluble CTLA4 muta	473	5	17.9	221	22	ABG05392	Novel human diagno
401	5	17.9	187	20	AAW41130	CTLA4 receptor wit	474	5	17.9	223	18	AAW25111	Soluble human CTLA
402	5	17.9	187	20	AAW97615	Human CTLA recepto	475	5	17.9	223	19	AAW98422	H. pylori GHPO 344
403	5	17.9	187	20	AAW97610	Human CTLA recepto	476	5	17.9	223	20	AAW41083	Canine CTLA4 prote
404	5	17.9	187	22	ABG13988	Novel human diagno	477	5	17.9	223	20	AAW41084	Feline CTLA4 prote
405	5	17.9	188	21	AAW13405	Arabidopsis thalia	478	5	17.9	223	21	AAW32287	Feline CTLA4-4. fe
406	5	17.9	188	21	AAW59844	Arabidopsis thalia	479	5	17.9	223	21	AAW15122	Porcine CTLA4-4 sol
407	5	17.9	188	21	AAW60284	Arabidopsis thalia	480	5	17.9	223	21	AAW15129	Human CTLA-4 prote
408	5	17.9	189	22	AAW82941	S. epidermidis ope	481	5	17.9	223	21	AAW32280	Cat CTLA-4 recepto
409	5	17.9	189	22	AAW82953	S. epidermidis ope	482	5	17.9	223	22	ABG12247	Human PI 5-phospha
410	5	17.9	191	21	AAW11917	S. epidermidis ope	483	5	17.9	223	22	AAW66519	Human CTLA4 Homo
411	5	17.9	191	21	AAW13404	Arabidopsis thalia	484	5	17.9	223	22	AAU00686	Porcine CTLA4 prot
412	5	17.9	191	21	AAW325102	Arabidopsis thalia	485	5	17.9	223	22	AAU00687	Human CTLA4 protei
413	5	17.9	191	21	AAW34702	Arabidopsis thalia	486	5	17.9	223	22	AAW62225	Glycine max glutat
414	5	17.9	193	21	AAW25105	Arabidopsis thalia	487	5	17.9	224	21	AAW17848	Arabidopsis thalia
415	5	17.9	193	21	AAW26364	Arabidopsis thalia	488	5	17.9	225	21	AAW17847	Arabidopsis thalia
416	5	17.9	194	22	ABW52721	Escherichia coli p	489	5	17.9	225	21	AAW20504	Arabidopsis thalia
417	5	17.9	195	19	AAW42428	Glucuronide repres	490	5	17.9	225	21	AAW47368	Arabidopsis thalia
418	5	17.9	195	21	AAW24926	Arabidopsis thalia	491	5	17.9	226	22	AAW79401	Human protein SEQ
419	5	17.9	195	21	AAW34466	Arabidopsis thalia	492	5	17.9	226	22	AAW94978	Human protein sequ
420	5	17.9	197	21	AAW13253	Caenorhabditis ele	493	5	17.9	228	12	AAW10743	Non-A non-B hepati
421	5	17.9	198	22	ABW67508	Drosophila melanog	494	5	17.9	228	21	AAW28350	Arabidopsis thalia
422	5	17.9	198	22	AAW91682	C glutamicum prote	495	5	17.9	228	21	AAW38010	Arabidopsis thalia
423	5	17.9	199	12	AAW10747	Non-A non-B hepati	496	5	17.9	228	21	AAW43642	Arabidopsis thalia
424	5	17.9	199	16	AAW73600	Frog TGF-beta 5 pr	497	5	17.9	228	22	AAW82384	S. epidermidis ope
425	5	17.9	199	21	AAW24925	Arabidopsis thalia	498	5	17.9	230	21	AAW24924	Arabidopsis thalia
426	5	17.9	199	22	ABW05697	Novel human diagno	499	5	17.9	230	21	AAW59843	Novel human diagno
427	5	17.9	200	22	AAW25823	Human protein sequ	500	5	17.9	230	22	ABG14565	Lung cancer associ
428	5	17.9	202	21	AAW11916	Arabidopsis thalia	501	5	17.9	231	21	AAW58136	Novel human diagno
429	5	17.9	202	21	AAW43701	Arabidopsis thalia	502	5	17.9	232	21	AAW21696	Arabidopsis thalia
430	5	17.9	204	13	AAW20986	Non-A, non-B hepat	503	5	17.9	232	22	AAU28373	Novel human secret
431	5	17.9	204	21	AAW17849	Arabidopsis thalia	504	5	17.9	233	21	AAW09629	Arabidopsis thalia
432	5	17.9	206	21	AAW42001	Human ORFX ORF1765	505	5	17.9	233	21	AAW61434	Arabidopsis thalia
433	5	17.9	206	21	AAW25101	Arabidopsis thalia	506	5	17.9	233	22	ABG01355	Novel human diagno
434	5	17.9	207	22	ABG16799	Novel human diagno	507	5	17.9	234	19	AAW61015	Streptococcus pneu
435	5	17.9	207	22	AAW79295	Amino acid sequenc	508	5	17.9	234	20	AAW41133	Propionibacterium
436	5	17.9	207	22	AAW95491	Human reproductive	509	5	17.9	234	20	AAW41134	Human breast cance
437	5	17.9	209	21	AAW25104	Arabidopsis thalia	510	5	17.9	236	21	AAW60283	Human neuronal gro
438	5	17.9	209	21	AAW94998	Human immune/haema	511	5	17.9	236	13	AAW25333	Sequence encoded b
439	5	17.9	209	22	ABG15735	Human CTLA4 recept	512	5	17.9	236	21	AAW81558	Human GAP-43. Hom
440	5	17.9	210	21	AAW41891	CTLA4 receptor fus	513	5	17.9	236	21	AAW81559	Human GAP-43. Hom
441	5	17.9	211	18	AAW14495	Full length CTLA4	514	5	17.9	237	22	AAU49141	Human GAP-43. Hom
442	5	17.9	211	20	AAW87560	Amino acid sequenc	515	5	17.9	237	22	AAW63234	Arabidopsis thalia
443	5	17.9	211	20	AAW87560	Human CTLA4 recept	516	5	17.9	238	11	AAW07038	Arabidopsis thalia
444	5	17.9	212	22	AAW91084	Human immune/haema	517	5	17.9	238	13	AAW25333	Streptococcus pneu
445	5	17.9	212	14	AAW31040	Human CTLA4 recept	518	5	17.9	238	15	AAW46486	Streptococcus pneu
446	5	17.9	212	15	AAW60134	CTLA4 receptor fus	519	5	17.9	238	15	AAW46497	Propionibacterium
447	5	17.9	212	16	AAW77642	Full length CTLA4	520	5	17.9	238	15	AAW46497	Human breast cance
448	5	17.9	212	20	AAW43479	Amino acid sequenc	521	5	17.9	238	21	AAW20141	Human neuronal gro
449	5	17.9	212	20	AAW81584	Human CTLA4 recept	522	5	17.9	238	21	AAW43740	Sequence encoded b
			212	21	AAW31659	Arabidopsis thalia						AAW48909	Human GAP-43. Hom

523	5	17.9	241	21	AAG35507	Arabidopsis thalia	596	5	17.9	286	16	AAR79659	P. aeruginosa tfda
524	5	17.9	241	21	AAG40180	Arabidopsis thalia	597	5	17.9	287	9	AAP80079	Sequence of 2,4-di
525	5	17.9	241	22	AAO12914	Human polypeptide	598	5	17.9	287	18	AAW31857	Mycobacterium tube
526	5	17.9	241	22	AAE03222	Human gene 4 encod	599	5	17.9	289	19	RAY10977	H. pylori ORF 02ge
527	5	17.9	242	22	AAE03205	Human gene 4 encod	600	5	17.9	290	21	AAV91644	Human secreted pro
528	5	17.9	243	19	AAW30561	Human regulator of	601	5	17.9	290	22	AAW25290	Human protein sequ
529	5	17.9	243	22	ABB68998	Drosophila melanog	602	5	17.9	291	18	AAW21018	H. pylori secreted
530	5	17.9	246	13	AAZ29893	HCV NS4-NS5 peptid	603	5	17.9	291	19	AAW98342	H. pylori GPHO 357
531	5	17.9	246	13	AAZ29894	HCV NS4-NS5 peptid	604	5	17.9	291	20	AAV34839	Chlamydia pneumoni
532	5	17.9	247	20	AAV48585	Human breast tumou	605	5	17.9	294	21	AAV35223	Zea mays protein f
533	5	17.9	248	20	AAV35735	Amino acid sequenc	606	5	17.9	295	22	AAU66743	Propionibacterium
534	5	17.9	248	21	AAG48901	Arabidopsis thalia	607	5	17.9	296	21	AAG12724	Arabidopsis thalia
535	5	17.9	248	22	AAW78417	Human protein SEQ	608	5	17.9	296	21	AAG52141	Arabidopsis thalia
536	5	17.9	249	22	AAU16412	Human novel secret	609	5	17.9	298	21	AAG28348	Arabidopsis thalia
537	5	17.9	250	21	AAV34702	Human secreted pro	610	5	17.9	298	21	AAG38008	Arabidopsis thalia
538	5	17.9	251	22	ABG21428	Novel human diagno	611	5	17.9	298	22	AAW40937	Human polypeptide
539	5	17.9	251	22	ABG25269	Novel human diagno	612	5	17.9	298	22	AAW40938	Human polypeptide
540	5	17.9	252	22	AAV51243	Human haemopoietin	613	5	17.9	298	22	AAW40939	Human polypeptide
541	5	17.9	253	21	AAV56947	Human prostate can	614	5	17.9	299	13	AAW25118	Non-A, Non-B Hepat
542	5	17.9	253	22	ABG15740	Novel human diagno	615	5	17.9	301	18	AAW23080	Metallophaepru
543	5	17.9	254	21	AAV61724	Arabidopsis thalia	616	5	17.9	301	21	AAG20502	Arabidopsis thalia
544	5	17.9	254	22	ABV70344	Drosophila melanog	617	5	17.9	301	21	AAG47366	Arabidopsis thalia
545	5	17.9	257	21	AAV25424	Pinus radiata cell	618	5	17.9	301	22	AAU40700	Propionibacterium
546	5	17.9	257	22	AAU46230	Propionibacterium	619	5	17.9	303	22	AAW25412	Human protein sequ
547	5	17.9	260	21	AAG28349	Arabidopsis thalia	620	5	17.9	304	21	AAG21902	Arabidopsis thalia
548	5	17.9	260	21	AAG38009	Arabidopsis thalia	621	5	17.9	305	21	AAV83390	Olfactory receptor
549	5	17.9	260	21	AAV95257	WOM-1 Fab heavy ch	622	5	17.9	305	22	AAU24533	Human olfactory re
550	5	17.9	260	22	AAU00685	Human CTLA4-human	623	5	17.9	305	22	AAE07545	Human G-protein co
551	5	17.9	261	22	AAU00684	Porcine CTLA4-huma	624	5	17.9	305	22	AAE07546	Human G-protein co
552	5	17.9	262	22	AAU54462	Propionibacterium	625	5	17.9	305	22	AAE07547	Human G-protein co
553	5	17.9	263	22	ABG10642	Novel human diagno	626	5	17.9	305	22	AAE07549	Human G-protein co
554	5	17.9	264	16	AAV78180	Rabbit alpha-d sub	627	5	17.9	305	22	AAV84951	Human olfactory re
555	5	17.9	264	18	AAW23063	Rabbit beta 2 inte	628	5	17.9	305	22	AAV71690	Human olfactory re
556	5	17.9	264	19	AAV72838	Rabbit alpha-d clo	629	5	17.9	305	22	AAV72596	Human OR-like poly
557	5	17.9	264	19	AAV5105	Rabbit beta-integr	630	5	17.9	306	21	AAG08377	Arabidopsis thalia
558	5	17.9	264	20	AAW60005	Rabbit alpha d pol	631	5	17.9	306	22	AAV31605	Arabidopsis thalia
559	5	17.9	264	20	AAW73349	Rabbit alpha d prot	632	5	17.9	306	22	ABG30265	Novel human diagno
560	5	17.9	264	21	AAV07375	Rabbit alpha d pro	633	5	17.9	307	22	ABG20066	Novel human diagno
561	5	17.9	264	22	ABG20639	Novel human diagno	634	5	17.9	308	18	AAW23084	Whale mat sample A
562	5	17.9	264	22	AAV98355	Escherichia coli p	635	5	17.9	312	21	AAV12723	Arabidopsis thalia
563	5	17.9	264	22	AAV981255	Human AFP protein	636	5	17.9	312	21	AAV25225	Arabidopsis thalia
564	5	17.9	265	21	AAV3641	Arabidopsis thalia	637	5	17.9	312	21	AAV99464	Arabidopsis thalia
565	5	17.9	265	22	AAV88462	Human membrane or	638	5	17.9	313	22	ABG12321	Novel human diagno
566	5	17.9	266	21	AAV52588	Helicobacter pylor	639	5	17.9	317	21	AAV17540	Arabidopsis thalia
567	5	17.9	266	22	AAV90461	C glutamicum prote	640	5	17.9	318	21	AAV29008	Arabidopsis thalia
568	5	17.9	268	19	AAV86041	S. pneumoniae deri	641	5	17.9	319	21	AAV17539	Arabidopsis thalia
569	5	17.9	268	22	AAV37659	Streptococcus pneu	642	5	17.9	319	22	AAV90187	C glutamicum prote
570	5	17.9	268	22	ABG17127	Novel human diagno	643	5	17.9	320	20	AAV41143	Mouse mammary tumo
571	5	17.9	269	21	AAV33293	Zea mays protein f	644	5	17.9	321	22	AAU17991	Human immunoglobul
572	5	17.9	269	21	AAV99445	Human PRO1787 (UNQ	645	5	17.9	322	19	AAV56249	Amino acid sequenc
573	5	17.9	269	21	AAV81604	Streptococcus pneu	646	5	17.9	322	20	AAV13379	Amino acid sequenc
574	5	17.9	269	22	AAU12436	Human PRO1787 poly	647	5	17.9	322	21	AAV87287	Human signal pepti
575	5	17.9	269	22	AAV66194	Protein of the inv	648	5	17.9	322	21	AAV87528	Human PRO263. Hom
576	5	17.9	271	21	AAV080036	Arabidopsis thalia	649	5	17.9	322	22	AAV88391	Human membrane or
577	5	17.9	272	13	AAV33884	NANBH fusion pepti	650	5	17.9	322	22	AAV80247	Human PRO263 prote
578	5	17.9	273	22	AAV93616	Human protein sequ	651	5	17.9	323	20	AAV88319	E. coli O157 antig
579	5	17.9	276	21	AAV30950	Arabidopsis thalia	652	5	17.9	323	22	ABV10354	Human cDNA seq ID
580	5	17.9	276	22	AAV94467	Human protein sequ	653	5	17.9	325	21	AAV20782	Arabidopsis thalia
581	5	17.9	277	18	AAV11700	Penicillium decumb	654	5	17.9	325	21	AAV40192	Arabidopsis thalia
582	5	17.9	277	21	AAV20503	Arabidopsis thalia	655	5	17.9	326	16	AAV85122	CDK6 R31C mutant.
583	5	17.9	277	21	AAV47367	Arabidopsis thalia	656	5	17.9	326	21	AAV35957	Zea mays protein f
584	5	17.9	277	22	AAV92526	C glutamicum prote	657	5	17.9	327	22	AAU35605	Haemophilus influe
585	5	17.9	277	22	AAV79026	Corynebacterium gl	658	5	17.9	328	21	AAV16348	E. grandis caffeic
586	5	17.9	277	22	AAV79937	Corynebacterium gl	659	5	17.9	329	22	AAV95056	Putative p. abyssi
587	5	17.9	279	21	AAV17541	Arabidopsis thalia	660	5	17.9	330	21	AAG12722	Arabidopsis thalia
588	5	17.9	280	22	AAV92705	Human protein sequ	661	5	17.9	331	22	AAG90146	C glutamicum prote
589	5	17.9	284	19	AAV61762	L. ivanovii strain	662	5	17.9	333	22	ABV58947	Drosophila melanog
590	5	17.9	284	19	AAV61763	L. seeligeri strai	663	5	17.9	333	22	ABV16843	Novel human diagno
591	5	17.9	284	19	AAV61764	L. welshimeri stra	664	5	17.9	334	13	AAV29878	HCV NS4-NS5 peptid
592	5	17.9	284	19	AAV61759	L. monocytogenes d	665	5	17.9	336	22	ABV21486	Novel human diagno
593	5	17.9	284	19	AAV61760	L. monocytogenes s	666	5	17.9	337	22	ABV70781	Drosophila melanog
594	5	17.9	284	19	AAV61761	L. innocua strain	667	5	17.9	337	22	AAU5094	Novel human diagno
595	5	17.9	285	21	AAV31606	Arabidopsis thalia	668	5	17.9	337	22	AAU35522	Human protein kina

669	5	17.9	338	21	AAG31323	Arabidopsis thalia	742	5	17.9	382	22	AAU42259	Propionibacterium
670	5	17.9	339	21	AAG35222	Zea mays protein f	743	5	17.9	383	21	AAB42481	Human ORFX ORF2245
671	5	17.9	339	21	AAV74561	Neisseria gonorrhoe	744	5	17.9	383	21	AAG58633	Arabidopsis thalia
672	5	17.9	339	21	AAV74562	Neisseria meningit	745	5	17.9	384	21	AAG51171	Arabidopsis thalia
673	5	17.9	339	21	AAV74563	Neisseria meningit	746	5	17.9	384	22	AAG89182	Human immune/haema
674	5	17.9	340	18	AAW33699	AL-2-short (AL-2s)	747	5	17.9	385	22	ABG02770	Novel human diagno
675	5	17.9	340	18	AAW31544	Human cytokine Ler	748	5	17.9	385	22	ABG02770	Human AFP protein
676	5	17.9	340	18	AAW10637	NLERK2 ligand for	749	5	17.9	387	21	AAV86261	Human secreted pro
677	5	17.9	340	18	AAW17081	EPH family ligand	750	5	17.9	388	22	AAE04798	Human gastricsin c
678	5	17.9	340	18	AAW46615	Human transmembran	751	5	17.9	388	22	AAE04798	Arabidopsis thalia
679	5	17.9	340	22	ABG08498	Novel human diagno	752	5	17.9	391	21	AAG45812	Zea mays protein f
680	5	17.9	340	22	ABG22055	Novel human diagno	753	5	17.9	395	12	AAG35221	Human AFP protein
681	5	17.9	344	18	AAW26767	Human chemokine re	754	5	17.9	395	12	AAG81359	R. communis delta9
682	5	17.9	344	19	AAW23957	Amino acid sequenc	755	5	17.9	396	12	AAAR14189	Ricinus communis d
683	5	17.9	344	22	ABG65507	Drosophila melanog	756	5	17.9	396	22	AAW44350	Novel human diagno
684	5	17.9	344	22	ABG11979	Human PRO263 homol	757	5	17.9	396	22	ABG07601	S. fradiae tylosin
685	5	17.9	345	12	AAAR12595	Antigenic portion	758	5	17.9	398	21	AAV83384	Human prostate can
686	5	17.9	345	21	AAG21695	Arabidopsis thalia	759	5	17.9	400	21	AAB56676	Arabidopsis thalia
687	5	17.9	347	22	AAG83116	S. epidermidis ope	760	5	17.9	400	21	AAG35816	Porcine CTLA-4-Ig
688	5	17.9	348	20	AAV37244	Amino acid sequenc	761	5	17.9	400	21	AAV15123	Novel human diagno
689	5	17.9	348	21	AAG06376	Arabidopsis thalia	762	5	17.9	401	19	AAW78920	Human haemochromat
690	5	17.9	348	21	AAG31604	Arabidopsis thalia	763	5	17.9	401	19	AAW69971	Human sodium-depen
691	5	17.9	351	17	AAAR87013	Xylanase A. Extre	764	5	17.9	403	21	AAG30833	Arabidopsis thalia
692	5	17.9	351	22	AAU15973	Human novel secret	765	5	17.9	403	21	ABY70199	Drosophila melanog
693	5	17.9	352	22	ABY58931	Drosophila melanog	766	5	17.9	404	22	AAV04663	L.lactis HsdS subu
694	5	17.9	353	21	AAG06375	Arabidopsis thalia	767	5	17.9	405	21	AAG58632	Arabidopsis thalia
695	5	17.9	353	22	ABY57766	Drosophila melanog	768	5	17.9	406	22	ABG19026	Novel human diagno
696	5	17.9	354	21	AAW58163	Lung cancer associ	769	5	17.9	407	21	AAG51170	Arabidopsis thalia
697	5	17.9	354	22	AAW73529	Human transferase	770	5	17.9	407	21	AAV77141	Native botulinum n
698	5	17.9	354	22	AAW74736	Human secreted pro	771	5	17.9	408	22	AAW38943	Human polypeptide
699	5	17.9	355	20	AAV02529	Amino acid sequenc	772	5	17.9	409	20	AAW94247	C. albicans tyrosy
700	5	17.9	356	19	AAW48087	Human macrophage/d	773	5	17.9	409	20	AAW94248	C. albicans tyrosy
701	5	17.9	356	21	AAV68798	Amino acid sequenc	774	5	17.9	409	21	AAV96727	Casein kinase II i
702	5	17.9	357	18	AAW35863	Human CTLA4:IG2a	775	5	17.9	409	22	ABY63209	Drosophila melanog
703	5	17.9	357	19	AAW68134	Amino acid sequenc	776	5	17.9	409	22	AAW51211	Human osteoclast e
704	5	17.9	357	21	AAW58365	Lung cancer associ	777	5	17.9	411	19	AAV37949	Amino acid sequenc
705	5	17.9	357	22	AAG71391	Human gene 22-enco	778	5	17.9	411	20	AAV48468	Human breast tumou
706	5	17.9	357	22	AAW94073	Human protein sequ	779	5	17.9	412	21	AAG45811	Arabidopsis thalia
707	5	17.9	359	21	AAG31322	Arabidopsis thalia	780	5	17.9	413	22	AAG84934	Botulism toxin hea
708	5	17.9	359	21	AAV93397	Amino acid sequenc	781	5	17.9	413	22	AAW04168	Sequence of a euka
709	5	17.9	360	13	AAAR29877	HCV NS4-NS5 peptid	782	5	17.9	414	14	AAAR32020	Human transcriptio
710	5	17.9	360	21	AAG25506	Arabidopsis thalia	783	5	17.9	414	19	AAW54406	Novel human diagno
711	5	17.9	360	21	AAW40179	Arabidopsis thalia	784	5	17.9	414	22	ABG08020	Human shear stress
712	5	17.9	363	22	AAW48199	Mature castor enzy	785	5	17.9	414	22	AAW90775	Putative P. abyss
713	5	17.9	364	21	AAW44797	Zea mays protein f	786	5	17.9	415	22	AAW96454	Drosophila melanog
714	5	17.9	364	21	AAV93698	Fusion protein of	787	5	17.9	416	22	ABY59251	Arabidopsis thalia
715	5	17.9	368	22	ABY63167	Drosophila melanog	788	5	17.9	417	21	AAW07679	Amino acid sequenc
716	5	17.9	368	22	AAW55176	Propionibacterium	789	5	17.9	417	21	AAG20781	Arabidopsis thalia
717	5	17.9	368	22	AAG81624	S. epidermidis ope	790	5	17.9	417	21	AAG40191	Arabidopsis thalia
718	5	17.9	368	22	AAW94843	Human protein sequ	791	5	17.9	417	22	ABG12620	Novel human diagno
719	5	17.9	370	22	ABY30007	Peptide #2658 enco	792	5	17.9	420	21	AAW83790	S. fradiae tylosin
720	5	17.9	370	22	ABY20615	Protein #2614 enco	793	5	17.9	420	22	AAU27554	Neisseria meningit
721	5	17.9	370	22	AAW56010	Human brain expres	794	5	17.9	421	22	AAW63222	Amino acid sequenc
722	5	17.9	370	22	AAW68380	Human bone marrow	795	5	17.9	423	21	AAB36302	C. botulinum BoNT/
723	5	17.9	370	22	AAW16198	Peptide #2632 enco	796	5	17.9	423	10	AAW90476	Polypeptide of hum
724	5	17.9	370	22	AAW03928	Peptide #2610 enco	797	5	17.9	429	21	AAG20780	Arabidopsis thalia
725	5	17.9	371	21	AAG14275	Arabidopsis thalia	798	5	17.9	429	21	AAG40190	Arabidopsis thalia
726	5	17.9	371	21	AAW46771	Arabidopsis thalia	799	5	17.9	430	22	ABG14501	Novel human diagno
727	5	17.9	371	22	ABG04190	Novel human diagno	800	5	17.9	432	21	AAV99620	Escherichia coli t
728	5	17.9	371	22	AAG51604	C glutamicum prote	801	5	17.9	432	21	ABG27575	E. coli trigger fa
729	5	17.9	372	21	AAG35818	Arabidopsis thalia	802	5	17.9	433	22	ABG27578	Novel human diagno
730	5	17.9	374	18	AAW26206	CTLA4-IgG4 fusion	803	5	17.9	433	22	AAW92838	C glutamicum prote
731	5	17.9	374	18	AAW26207	CTLA4-modified IgG	804	5	17.9	435	20	AAV37410	Protein involved i
732	5	17.9	374	21	AAG35817	Arabidopsis thalia	805	5	17.9	435	21	AAW20519	Agrocycbes pedabes
733	5	17.9	376	21	AAW86190	Nuclear transport	806	5	17.9	435	21	AAW69562	Agrocycbe pedabes
734	5	17.9	376	22	ABG15410	Novel human diagno	807	5	17.9	436	22	ABY58109	Drosophila melanog
735	5	17.9	377	18	AAW26208	CTLA4-modified IgG	808	5	17.9	438	20	AAW95500	B. subtilis YabE r
736	5	17.9	380	12	AAAR10748	Non-A non-B hepati	809	5	17.9	440	21	AAG35956	Zea mays protein f
737	5	17.9	380	21	AAG30834	Arabidopsis thalia	810	5	17.9	440	21	AAV75938	Neisseria strain z
738	5	17.9	381	19	AAW81597	Protein encoded by	811	5	17.9	440	22	ABG27572	Novel human diagno
739	5	17.9	381	22	AAW74765	Human secreted pro	812	5	17.9	440	22	AAU27555	Neisseria meningit
740	5	17.9	382	21	AAW08338	Amino acid sequenc	813	5	17.9	441	15	AAAR43992	Bovine ET receptor
741	5	17.9	382	21	AAG17082	Arabidopsis thalia	814	5	17.9	441	21	AAW58565	N. meningitidis am

815	5	17.9	441	21	AAB58566	N. meningitidis am	888	5	17.9	454	22	AAU36519	Pseudomonas aerugi
816	5	17.9	441	21	AAB58567	N. gonorrhoea amin	889	5	17.9	455	18	AAW33698	AL-2-long (AL-21)
817	5	17.9	441	21	AAB58568	N. meningitidis am	890	5	17.9	455	22	ABB11613	Human sorting nexi
818	5	17.9	441	21	AAB58569	N. meningitidis am	891	5	17.9	456	16	AAW49996	E. maxima Em70-1 a
819	5	17.9	441	21	AAB58570	N. meningitidis am	892	5	17.9	456	22	AAU35711	Helicobacter pylori
820	5	17.9	441	21	AAB58571	N. meningitidis am	893	5	17.9	457	21	AAW33321	Arabidopsis thalia
821	5	17.9	441	21	AAB58572	N. meningitidis am	894	5	17.9	457	22	AAW41619	Human polypeptide
822	5	17.9	441	21	AAB58573	N. meningitidis am	895	5	17.9	458	22	AAW96675	Putative P. abyssi
823	5	17.9	441	21	AAB58574	Neisseria gonorrhoe	896	5	17.9	458	22	AAW47297	PRO9912 polypeptid
824	5	17.9	441	21	AAB58575	Neisseria meningit	897	5	17.9	459	21	AAW30832	Arabidopsis thalia
825	5	17.9	441	21	AAB58576	Neisseria meningit	898	5	17.9	461	22	AAW91287	C glutamicum prote
826	5	17.9	441	21	AAB58577	Neisseria meningit	899	5	17.9	463	19	AAW48761	Human zinc finger
827	5	17.9	441	21	AAB58578	Neisseria meningit	900	5	17.9	463	21	AAW35955	zebra mays protein f
828	5	17.9	441	21	AAB58579	Neisseria strain f	901	5	17.9	463	21	AAW51169	Arabidopsis thalia
829	5	17.9	441	21	AAB58580	Neisseria strain z	902	5	17.9	464	22	AAW88343	Human membrane or
830	5	17.9	441	21	AAB58581	Neisseria strain z	903	5	17.9	466	22	AAW38424	Salmonella typhi c
831	5	17.9	441	21	AAB58582	Neisseria strain z	904	5	17.9	468	21	AAW45810	Arabidopsis thalia
832	5	17.9	441	21	AAB58583	Neisseria strain z	905	5	17.9	469	21	AAW43932	Human cancer assoc
833	5	17.9	441	21	AAB58584	Neisseria strain z	906	5	17.9	469	22	AAW39153	Human polypeptide
834	5	17.9	441	21	AAB58585	Neisseria strain z	907	5	17.9	470	19	AAW68135	Amino acid sequenc
835	5	17.9	441	21	AAB58586	Neisseria strain z	908	5	17.9	470	20	AAW17215	H. pylori outer me
836	5	17.9	441	21	AAB58587	Neisseria strain z	909	5	17.9	472	20	AAW60508	Human normal blad
837	5	17.9	441	21	AAB58588	Neisseria strain z	910	5	17.9	476	20	AAW59725	Human normal ovari
838	5	17.9	441	21	AAB58589	Neisseria strain z	911	5	17.9	478	22	AAW18941	Novel human diagno
839	5	17.9	441	21	AAB58590	Neisseria strain z	912	5	17.9	479	20	AAW94283	Human protein phos
840	5	17.9	441	21	AAB58591	Neisseria strain z	913	5	17.9	479	21	AAW42154	Human ORFX ORF1918
841	5	17.9	441	21	AAB58592	Neisseria strain z	914	5	17.9	479	21	AAW14274	Arabidopsis thalia
842	5	17.9	441	21	AAB58593	Neisseria strain z	915	5	17.9	479	21	AAW46770	Arabidopsis thalia
843	5	17.9	441	21	AAB58594	Neisseria strain z	916	5	17.9	480	19	AAW98219	H. pylori GHPO 170
844	5	17.9	441	21	AAB58595	Neisseria strain z	917	5	17.9	482	21	AAW32191	Arabidopsis thalia
845	5	17.9	441	21	AAB58596	Neisseria strain z	918	5	17.9	483	15	AAW58590	Hepatitis C virus
846	5	17.9	441	21	AAB58597	Neisseria strain z	919	5	17.9	484	22	AAW34864	E. coli cellular p
847	5	17.9	441	21	AAB58598	Neisseria strain z	920	5	17.9	485	22	ABG15750	Novel human diagno
848	5	17.9	441	21	AAB58599	Neisseria strain z	921	5	17.9	486	21	AAW14273	Arabidopsis thalia
849	5	17.9	441	21	AAB58600	Neisseria strain z	922	5	17.9	486	21	AAW46769	Arabidopsis thalia
850	5	17.9	441	21	AAB58601	Neisseria strain z	923	5	17.9	486	22	ABW70458	Drosophila melanog
851	5	17.9	441	21	AAB58602	Neisseria strain z	924	5	17.9	486	22	AAU37692	Streptococcus pneu
852	5	17.9	441	21	AAB58603	Neisseria strain z	925	5	17.9	486	22	AAU37692	CFE 6 protein sequ
853	5	17.9	441	21	AAB58604	Neisseria strain z	926	5	17.9	489	22	AAW01007	Drosophila melanog
854	5	17.9	441	21	AAB58605	Neisseria strain z	927	5	17.9	491	22	AAW83319	M. sativa PDH prot
855	5	17.9	441	21	AAB58606	Neisseria strain z	928	5	17.9	491	22	AAW93582	Human protein sequ
856	5	17.9	441	21	AAB58607	Neisseria strain z	929	5	17.9	492	21	AAW43110	Human ORFX ORF2874
857	5	17.9	441	21	AAB58608	Neisseria strain z	930	5	17.9	493	21	AAW41564	Human ORFX ORF1328
858	5	17.9	441	21	AAB58609	Neisseria strain z	931	5	17.9	493	22	ABG29241	Novel human diagno
859	5	17.9	441	21	AAB58610	Neisseria strain z	932	5	17.9	496	22	ABW12018	Human acetylcholin
860	5	17.9	441	21	AAB58611	Neisseria strain z	933	5	17.9	496	22	AAW73042	Computer-generated
861	5	17.9	441	21	AAB58612	Neisseria strain z	934	5	17.9	497	20	AAW74187	Human prostate tum
862	5	17.9	441	21	AAB58613	Neisseria meningit	935	5	17.9	497	22	AAW90137	C glutamicum prote
863	5	17.9	441	21	AAB58614	Amino acid sequenc	936	5	17.9	501	22	AAW96488	Putative P. abyssi
864	5	17.9	441	21	AAB58615	Neisseria MenB 919	937	5	17.9	502	22	ABW71611	Drosophila melanog
865	5	17.9	441	21	AAB58616	Human polypeptide	938	5	17.9	502	22	ABG07742	Novel human diagno
866	5	17.9	441	21	AAB58617	Human protein havi	939	5	17.9	508	21	AAW32190	Arabidopsis thalia
867	5	17.9	441	21	AAB58618	Peptide #3486 enco	940	5	17.9	510	22	AAW93365	Human protein sequ
868	5	17.9	441	21	AAB58619	Peptide #3644 enco	941	5	17.9	510	22	AAU00471	Human TANGO 364 pr
869	5	17.9	441	21	AAB58620	Peptide #3518 enco	942	5	17.9	515	22	AAU60003	Propionibacterium
870	5	17.9	441	21	AAB58621	Peptide #3692 enco	943	5	17.9	518	22	AAU28185	Novel human secret
871	5	17.9	441	21	AAB58622	Human bone marrow	944	5	17.9	518	22	AAW92833	Human protein sequ
872	5	17.9	441	21	AAB58623	Human bone marrow	945	5	17.9	518	22	AAW93296	Human Voltage-gate
873	5	17.9	441	21	AAB58624	Peptide #3453 enco	946	5	17.9	519	21	AAW44565	Human Voltage-gate
874	5	17.9	441	21	AAB58625	Peptide #3619 enco	947	5	17.9	519	21	AAW44569	Human Voltage-gate
875	5	17.9	441	21	AAB58626	Peptide #3548 enco	948	5	17.9	519	21	AAW44570	Human Voltage-gate
876	5	17.9	441	21	AAB58627	Peptide #3715 enco	949	5	17.9	519	21	AAW44571	Human Voltage-gate
877	5	17.9	441	21	AAB58628	Peptide #3408 enco	950	5	17.9	519	21	AAW44572	Human Voltage-gate
878	5	17.9	441	21	AAB58629	Peptide #3562 enco	951	5	17.9	523	14	AAW37493	cdc25A. Homo sapi
879	5	17.9	441	21	AAB58630	C. elegans delta 5	952	5	17.9	523	16	AAW66608	Human cdc25A gene
880	5	17.9	441	21	AAB58631	Human fatty a	953	5	17.9	523	17	AAW98212	Human CDC25A phosph
881	5	17.9	441	21	AAB58632	Human colon cancer	954	5	17.9	523	19	AAW37986	Amino acid sequenc
882	5	17.9	441	21	AAB58633	Hepatitis C virus	955	5	17.9	523	19	AAW59135	Human cdc25A prote
883	5	17.9	441	21	AAB58634	UTR1 gene product.	956	5	17.9	523	21	AAW54905	Human cdc25A prote
884	5	17.9	441	21	AAB58635	Acrocybe pedicels.	957	5	17.9	523	22	ABG15092	Novel human diagno
885	5	17.9	441	21	AAB58636	A. pedicels phytas	958	5	17.9	523	22	AAW68001	Amino acid sequenc
886	5	17.9	441	21	AAB58637	Protein 25bee with	959	5	17.9	523	22	AAW81084	Human Cdc25A prote
887	5	17.9	441	21	AAB58638	Drosophila melanog	960	5	17.9	523	22	AAW20263	Human Cdc25 phosph

961	5	17.9	524	22	ABG14128	Novel human diagno	1034	5	17.9	652	22	AAB51242	Human haemopoietin
962	5	17.9	528	19	AAW69963	Aeromonas caviae x	1035	5	17.9	655	16	AAR74044	Carbonic-anhydrase
963	5	17.9	529	22	AAW01071	CFe 74 protein seq	1036	5	17.9	656	22	ABG11615	Novel human diagno
964	5	17.9	529	22	ABG95631	Human protein sequ	1037	5	17.9	656	22	AAB92919	Human protein sequ
965	5	17.9	532	22	ABG20407	Novel human diagno	1038	5	17.9	658	21	AAG36831	Arabidopsis thalia
966	5	17.9	536	16	AAR76502	Human MIS mature p	1039	5	17.9	662	22	AAB51244	Human haemopoietin
967	5	17.9	536	20	AAW73937	Human prostate tum	1040	5	17.9	663	21	AAB18996	A partial core 2 b
968	5	17.9	536	22	ABG15944	Novel human diagno	1041	5	17.9	665	22	AAM39430	Human polypeptide
969	5	17.9	540	22	AAE05683	Brassica napus cyt	1042	5	17.9	666	19	AAW72911	Mycobacterium tube
970	5	17.9	540	22	AAR73536	Streptococcus pneu	1043	5	17.9	666	20	AAW21928	Amino acid sequenc
971	5	17.9	541	22	ABG60815	Drosophila melanog	1044	5	17.9	668	22	AAB19846	Mycobacterium tube
972	5	17.9	542	19	AAR79072	Sinapis alba cytoc	1045	5	17.9	668	21	AAB41651	Human ORFX ORF1415
973	5	17.9	542	20	AAW35078	Chlamydia pneumoni	1046	5	17.9	668	22	AAB94215	Human protein sequ
974	5	17.9	546	16	AAR74045	Carbonic-anhydrase	1047	5	17.9	669	22	ABG11302	Drosophila melanog
975	5	17.9	547	22	ABG65090	Drosophila melanog	1048	5	17.9	672	21	AAB36170	Novel human transp
976	5	17.9	547	22	ABG67062	Drosophila melanog	1049	5	17.9	672	21	AAB36171	Novel human transp
977	5	17.9	549	22	ABG09715	Novel human diagno	1050	5	17.9	684	22	ABG11697	Novel human diagno
978	5	17.9	549	22	AAU01899	Mycobacterium tube	1051	5	17.9	684	22	ABG12169	Novel human diagno
979	5	17.9	551	21	AAB42049	Human ORFX ORF1813	1052	5	17.9	685	22	ABG11107	Drosophila melanog
980	5	17.9	556	20	AAW35178	Chlamydia pneumoni	1053	5	17.9	685	22	ABG11254	Drosophila melanog
981	5	17.9	556	21	AAB18820	A 60 kDa cysteine-	1054	5	17.9	689	14	AAR30509	N-terminal of LH r
982	5	17.9	559	8	AAW70196	Sequence cystede-b	1055	5	17.9	689	22	AAB68912	Drosophila melanog
983	5	17.9	560	16	AAR76501	Human MIS protein.	1056	5	17.9	689	22	AAB38933	Human polypeptide
984	5	17.9	560	21	AAW92020	Human Mullerian In	1057	5	17.9	690	21	AAG23045	Arabidopsis thalia
985	5	17.9	561	22	ABG07602	Novel human diagno	1058	5	17.9	691	21	AAG23044	Arabidopsis thalia
986	5	17.9	568	22	ABG68716	Drosophila melanog	1059	5	17.9	691	22	ABG03068	Human expressed po
987	5	17.9	568	22	AAW33836	Human polypeptide,	1060	5	17.9	691	22	AAU23180	Novel human enzyme
988	5	17.9	569	22	ABG21491	Novel human diagno	1061	5	17.9	692	11	AAR08038	Rat testicular lut
989	5	17.9	570	22	ABG20901	Novel human diagno	1062	5	17.9	692	14	AAR30503	N-terminal of LH r
990	5	17.9	575	21	AAB42376	Human ORFX ORF2140	1063	5	17.9	692	14	AAR30506	N-terminal of LH r
991	5	17.9	576	21	AAB34234	Human vesicle asso	1064	5	17.9	695	14	AAR30524	N-terminal of LH r
992	5	17.9	576	22	AAW95100	Human protein sequ	1065	5	17.9	695	14	AAR30525	N-terminal of LH r
993	5	17.9	584	22	ABG82869	S. epidermidis ope	1066	5	17.9	695	22	ABG62607	Drosophila melanog
994	5	17.9	586	22	ABG28222	Drosophila melanog	1067	5	17.9	695	14	AAR30523	N-terminal of LH r
995	5	17.9	586	22	ABG67074	Drosophila melanog	1068	5	17.9	697	21	AAG42037	Arabidopsis thalia
996	5	17.9	587	18	AAW33274	S. fradiae tylosin	1069	5	17.9	698	14	AAR30505	N-terminal of LH r
997	5	17.9	589	22	ABG16367	Novel human diagno	1070	5	17.9	698	22	AAM38932	Human polypeptide
998	5	17.9	590	12	AAW12594	Antigenic portion	1071	5	17.9	699	14	AAR30517	N-terminal of LH r
999	5	17.9	596	14	AAW14433	PT-NANBH virus NS5	1072	5	17.9	699	14	AAR30516	N-terminal of LH r
1000	5	17.9	597	22	ABG60160	Drosophila melanog	1073	5	17.9	700	14	AAR30511	N-terminal of LH r
1001	5	17.9	598	21	AAW91963	PMWav-1 helicase.	1074	5	17.9	706	22	AAG62870	Amino acid sequenc
1002	5	17.9	601	22	ABG01164	Novel human diagno	1075	5	17.9	706	22	AAG60648	Human protein kina
1003	5	17.9	603	22	AAU04485	Human protein tyro	1076	5	17.9	707	22	AAB65526	Drosophila melanog
1004	5	17.9	605	20	AAW05697	Human TNF receptor	1077	5	17.9	709	21	AAW93338	A murine interphot
1005	5	17.9	606	13	AAW29880	HCV NS4-NS5 peptid	1078	5	17.9	709	22	AAM93837	Human polypeptide,
1006	5	17.9	606	21	AAB42236	Human ORFX ORF2000	1079	5	17.9	712	22	ABG27220	Novel human diagno
1007	5	17.9	609	14	AAR32131	CMV Colburn region	1080	5	17.9	717	22	ABG24062	Novel human diagno
1008	5	17.9	611	22	AAW40718	Human polypeptide	1081	5	17.9	718	22	ABG03655	Novel human diagno
1009	5	17.9	611	22	AAW40719	Human polypeptide	1082	5	17.9	723	22	ABG06215	Novel human diagno
1010	5	17.9	612	22	ABW58181	Drosophila melanog	1083	5	17.9	727	18	AAW11719	C-Delta-1 polypept
1011	5	17.9	613	22	AAW95862	Human protein sequ	1084	5	17.9	728	21	AAW79029	Chick delta protei
1012	5	17.9	618	22	ABW66120	Drosophila melanog	1085	5	17.9	728	22	ABW11763	Drosophila melanog
1013	5	17.9	620	19	AAW50907	Alteromonas carra	1086	5	17.9	732	22	AAU29362	Novel mar regulate
1014	5	17.9	622	21	AAG32189	Arabidopsis thalia	1087	5	17.9	735	21	AAW74556	Neisseria meningit
1015	5	17.9	622	22	ABW94133	Human protein sequ	1088	5	17.9	740	18	AAW00876	C-Delta-1 polypept
1016	5	17.9	627	22	ABW68963	Drosophila melanog	1089	5	17.9	740	22	AAB94340	Human protein sequ
1017	5	17.9	627	22	ABG29833	Novel human diagno	1090	5	17.9	744	22	ABG02772	Novel human diagno
1018	5	17.9	629	21	AAG48916	Arabidopsis thalia	1091	5	17.9	746	21	AAW44358	P. chrysogenum suc
1019	5	17.9	630	18	AAR32213	Mutase of penici	1092	5	17.9	748	21	AAG42036	Arabidopsis thalia
1020	5	17.9	630	20	AAW97037	Penicillium purpur	1093	5	17.9	748	22	ABW59560	Drosophila melanog
1021	5	17.9	630	20	AAW89392	Penicillium purpur	1094	5	17.9	758	21	AAB18719	A Neisseria mening
1022	5	17.9	631	22	ABW65629	Drosophila melanog	1095	5	17.9	758	21	AAW74564	Neisseria gonorrhe
1023	5	17.9	633	22	AAW96656	Putative P. abyss	1096	5	17.9	758	21	AAW74565	Neisseria meningit
1024	5	17.9	634	14	AAR30520	N-terminal of LH r	1097	5	17.9	758	21	AAW74566	Neisseria meningit
1025	5	17.9	639	21	AAG48915	Arabidopsis thalia	1098	5	17.9	758	22	AAG28071	Novel human secret
1026	5	17.9	645	21	AAG48914	Arabidopsis thalia	1099	5	17.9	761	22	ABG15896	Novel human diagno
1027	5	17.9	645	22	AAG82945	S. epidermidis ope	1100	5	17.9	761	22	AAE09794	Soybean oxidosqual
1028	5	17.9	648	20	AAW52834	Rhodobacter sphaer	1101	5	17.9	762	22	ABW67496	Drosophila melanog
1029	5	17.9	650	20	AAW14962	Amino acid sequenc	1102	5	17.9	762	22	AAE09795	Soybean beta-amylri
1030	5	17.9	650	22	AAB83262	C elegans FATPa SE	1103	5	17.9	764	21	AAW74557	Neisseria meningit
1031	5	17.9	650	22	AAB83274	C elegans FATPa SE	1104	5	17.9	766	18	AAW37352	PsBC (WbpC) protei
1032	5	17.9	651	16	AAR74042	Maize carbonic-ahn	1105	5	17.9	766	22	ABG14630	Novel human diagno
1033	5	17.9	652	21	AAB42426	Human ORFX ORF2190	1106	5	17.9	767	22	AAM41216	Human polypeptide

1107	5	17.9	774	21	AAG36830	Arabidopsis thalia
1108	5	17.9	779	22	ABG06522	Novel human diagno
1109	5	17.9	783	22	ABB37617	Peptide #5123 enco
1110	5	17.9	783	22	ABB22911	Protein #4910 enco
1111	5	17.9	783	22	AAM58270	Human brain expres
1112	5	17.9	783	22	AAM70723	Human bone marrow
1113	5	17.9	783	22	AAM18576	Peptide #5010 enco
1114	5	17.9	783	22	AAM31036	Peptide #5073 enco
1115	5	17.9	785	21	AAG42035	Arabidopsis thalia
1116	5	17.9	794	22	AGE84963	Shrimp white spot
1117	5	17.9	798	22	AAE04364	Human kinase (PKIN
1118	5	17.9	802	22	ABG27812	Novel human diagno
1119	5	17.9	803	13	AAR29030	Bacillus thuringie
1120	5	17.9	803	13	AAR28814	BT toxin 63B. Bac
1121	5	17.9	803	13	AAR29521	BT toxin 63B. Bac
1122	5	17.9	803	18	AAW13887	63B toxin. Bacill
1123	5	17.9	803	21	AAW13913	Bacillus thuringie
1124	5	17.9	803	22	ABG59885	63B protein. Baci
1125	5	17.9	804	21	AAW94905	Human secreted pro
1126	5	17.9	805	22	ABG66991	Drosophila melanog
1127	5	17.9	808	22	AAW94734	Human protein sequ
1128	5	17.9	811	21	ABG36829	Arabidopsis thalia
1129	5	17.9	811	22	AAE06027	Human asthma-assoc
1130	5	17.9	812	22	ABG66249	Drosophila melanog
1131	5	17.9	813	20	AAW93598	O. sativa Xa21 gen
1132	5	17.9	814	22	ABB71551	Drosophila melanog
1133	5	17.9	816	16	AAR85870	WD-40 domain-contg
1134	5	17.9	819	22	ABG4944	Drosophila melanog
1135	5	17.9	822	22	ABG57937	Drosophila melanog
1136	5	17.9	824	22	ABG61091	Drosophila melanog
1137	5	17.9	824	22	ABG61100	Drosophila melanog
1138	5	17.9	826	21	AAW07228	Human prostate can
1139	5	17.9	826	21	AAW99850	Human sulphatase G
1140	5	17.9	832	22	AAW62517	Drosophila melanog
1141	5	17.9	832	22	AAU27569	Neisseria meningit
1142	5	17.9	832	22	AAE10014	N. meningitidis st
1143	5	17.9	835	22	AAW35401	Human PG-3. Homo
1144	5	17.9	837	21	AAW77140	Native botulinum n
1145	5	17.9	847	22	AAW04081	Botulism toxin hea
1146	5	17.9	853	21	AAW26917	Large subunit of p
1147	5	17.9	854	22	AAW32918	Human protein sequ
1148	5	17.9	863	13	AAR29881	HCV NS4-NS5 peptid
1149	5	17.9	863	22	ABG90822	Human shear stress
1150	5	17.9	864	22	ABG27219	Novel human diagno
1151	5	17.9	868	22	ABG24569	Novel human diagno
1152	5	17.9	871	19	AAW56019	Recombinant botuli
1153	5	17.9	871	19	AAW56007	Recombinant botuli
1154	5	17.9	871	19	AAW56008	Botulinum neurotox
1155	5	17.9	873	19	AAW56016	Recombinant botuli
1156	5	17.9	874	22	AAW95467	Human protein sequ
1157	5	17.9	875	19	AAW56009	Recombinant botuli
1158	5	17.9	876	20	AAW41293	Human soluble prot
1159	5	17.9	876	22	ABG36597	Human FLEXHT-19 pr
1160	5	17.9	877	22	ABG28379	Novel human diagno
1161	5	17.9	878	16	AAR78608	Murine IL-3 recept
1162	5	17.9	878	17	AAR92529	Fas sequence from
1163	5	17.9	878	19	AAW56010	Recombinant botuli
1164	5	17.9	880	22	AAW96332	Putative P. abyssi
1165	5	17.9	880	22	AAW70938	S. pombe potassium
1166	5	17.9	884	22	ABG63254	Drosophila melanog
1167	5	17.9	887	20	AAW15344	Tumour suppressor
1168	5	17.9	887	20	AAW28995	Tumour suppressor
1169	5	17.9	889	22	ABG66958	Drosophila melanog
1170	5	17.9	890	22	ABW1207	Drosophila melanog
1171	5	17.9	892	22	AAU38250	Salmonella typhi c
1172	5	17.9	893	22	ABG63245	Drosophila melanog
1173	5	17.9	894	19	AAW56015	Recombinant botuli
1174	5	17.9	897	22	AAU27572	Neisseria meningit
1175	5	17.9	897	22	AAE10017	N. meningitidis st
1176	5	17.9	898	22	ABG12152	Novel human diagno
1177	5	17.9	899	22	AAE10153	Protein encoded by
1178	5	17.9	900	22	ABG62018	Drosophila melanog
1179	5	17.9	901	22	AAW95108	Human protein sequ

RESULT 1

AAW03189	1180	5	17.9	903	20	AAW99665	Human secreted pro
ID	1181	5	17.9	903	21	AAG44774	Arabidopsis thalia
XX	1182	5	17.9	907	19	AAW56012	Recombinant botuli
AC	1183	5	17.9	908	22	ABB11869	Human soluble prot
XX	1184	5	17.9	917	16	AAR80132	Rps2 disease resis
DT	1185	5	17.9	919	22	ABG06857	Novel human diagno
XX	1186	5	17.9	921	15	AAW56487	TATA-binding prote
XX	1187	5	17.9	921	17	AAW06077	Drosophila TATA-bi
DE	1188	5	17.9	921	18	AAW25028	TATA-binding prote
XX	1189	5	17.9	921	22	ABB61528	Drosophila melanog
XX	1190	5	17.9	921	22	ABB66055	Drosophila melanog
XX	1191	5	17.9	930	14	AAR41431	PT-NANBH virus BHC
XX	1192	5	17.9	939	22	ABB58143	Drosophila melanog
XX	1193	5	17.9	944	22	ABB61800	Drosophila melanog
XX	1194	5	17.9	949	22	ABG06446	Novel human diagno
XX	1195	5	17.9	950	22	ABB62035	Drosophila melanog
XX	1196	5	17.9	953	19	AAW56011	Recombinant botuli
XX	1197	5	17.9	956	14	AAR32356	Excitatory amino aci
XX	1198	5	17.9	956	22	ABG12450	Novel human diagno
XX	1199	5	17.9	957	22	ABG26105	Novel human diagno
XX	1200	5	17.9	957	22	AAU07894	Polypeptide sequen

ALIGNMENTS

AAW03189 standard; peptide; 20 AA.

23-OCT-2000 (first entry)

Human selenoprotein P C-terminal fragment, SEQ ID NO:7.
 Selenoprotein P fragment; human; programmed cell death inhibitor;
 apoptosis inhibitor; antiapoptotic; acquired immunodeficiency syndrome;
 AIDS; Parkinson's disease; Alzheimer's disease; myocardial infarct;
 cerebral infarct; stroke; reperfusion injury; organ transplant rejection;
 arteriosclerosis; redox disorder; immune disorder; cardiovascular;
 cerebroprotective; culture medium additive.

Key	Location/Qualifiers
Modified-site	/note= "Selenocysteine"
Modified-site	19
Modified-site	/note= "Selenocysteine"
WO200031131-A1.	

02-JUN-2000.

12-NOV-1999; 99WO-JP063322.

19-NOV-1998; 98JP-0347863.

(KAGA) CHEMO-SERO-THERAPEUTIC RES INST.

Hirashima M, Maeda H, Nozaki C;

WPI; 2000-451677/39.

Peptide fragments with cell death inhibitory activity, useful in preventing and treating apoptosis-associated diseases particularly caused by stress e.g. Parkinson's disease, Alzheimer's and arteriosclerosis

Example 7; Page 25; 56pp; Japanese.

The invention relates to fragments (AAB03183, AAB03184) derived from the

CC C-terminal 103 amino acids of human selenoprotein P (AAB03188) which act
 CC as inhibitors of apoptosis (programmed cell death), their use in
 CC therapeutic compositions and cell culture media, and antibodies against
 CC them. The invention also relates to a method for studying the phenomenon
 CC of sudden cell death in a human megakaryotic series gemmule cell culture
 CC system involving use of the anti-apoptotic peptides of the invention.
 CC Selenoprotein P is a plasma protein which confers protection against
 CC peroxynitrite-mediated oxidation and nitration. The peptide fragments
 CC can be used as prophylactic and therapeutic agents for
 CC apoptosis-associated diseases, including AIDS (acquired immunodeficiency
 CC syndrome), Parkinson's disease, Alzheimer's disease, arteriosclerosis,
 CC myocardial infarct, cerebral infarct, organ transplant rejection, and
 CC reperfusion injury. They may also be used in the prevention or treatment
 CC of disorders relating to redox reactions, or of disorders of the immune
 CC system. They may additionally be used as an additive for cell culture and
 CC for screening cell death activity in vitro. The present sequence
 CC represents a C-terminal fragment of human selenoprotein P used in an
 CC exemplification of the invention.

SQ Sequence 20 AA;

Query Match 39.3%; Score 11; DB 21; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 QCKENLPSPSICS 18
 | | | | | | | | | |
 Db 8 qckenlpsslcs 18

RESULT 2

AAB03184
 ID AAB03184 standard; peptide; 28 AA.

AC AAB03184;

XX 23-OCT-2000 (first entry)

XX Human selenoprotein P C-terminal fragment, SEQ ID NO:2.

XX Selenoprotein P fragment; human; programmed cell death inhibitor;
 KW apoptosis inhibitor; antiapoptotic; acquired immunodeficiency syndrome;
 KW AIDS; Parkinson's disease; Alzheimer's disease; myocardial infarct;
 KW cerebral infarct; stroke; reperfusion injury; organ transplant rejection;
 KW arteriosclerosis; redox disorder; immune disorder; cardiovascular;
 KW cerebroprotective; culture medium additive.

XX Homo sapiens.

Key	Location/Qualifiers
Modified-site 7	/note= "Selenocysteine"
Modified-site 19	/note= "Selenocysteine"

XX W0200031131-AL.

XX 02-JUN-2000.

XX 12-NOV-1999; 99WO-JP06322.

XX 19-NOV-1998; 98JP-0347863.

XX (KAGA) CHEMO-SERO-THERAPEUTIC RES INST.

XX Hirashima M, Maeda H, Nozaki C;

XX WPI; 2000-451677/39.

XX Peptide fragments with cell death inhibitory activity, useful in
 PT preventing and treating apoptosis-associated diseases particularly
 PT caused by stress e.g. Parkinson's disease, Alzheimer's and

PT arteriosclerosis -
 XX Claim 2; Page 31; 56pp; Japanese.
 XX This sequence represents a fragment of human selenoprotein P which
 CC has anti-apoptotic activity. The invention relates to fragments
 CC (AAB03183, AAB03184) derived from the C-terminal 103 amino acids of
 CC human selenoprotein P (AAB03188) which act as inhibitors of apoptosis
 CC (programmed cell death), their use in therapeutic compositions and cell
 CC culture media, and antibodies against them. The invention also relates to
 CC a method for studying the phenomenon of sudden cell death in a human
 CC megakaryotic series gemmule cell culture system involving use of the
 CC anti-apoptotic peptides of the invention. Selenoprotein P is a plasma
 CC protein which confers protection against peroxynitrite-mediated
 CC oxidation and nitration. The peptide fragments can be used as
 CC prophylactic and therapeutic agents for apoptosis-associated diseases,
 CC including AIDS (acquired immunodeficiency syndrome), Parkinson's disease,
 CC Alzheimer's disease, arteriosclerosis, myocardial infarct, cerebral
 CC infarct, organ transplant rejection, and reperfusion injury. They may
 CC also be used in the prevention or treatment of disorders relating to
 CC redox reactions, or of disorders of the immune system. They may
 CC additionally be used as an additive for cell culture and for screening
 CC cell death activity in vitro.

SQ Sequence 28 AA;

Query Match 39.3%; Score 11; DB 21; Length 28;

Best Local Similarity 100.0%; Pred. No. 6.4e-05;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 QCKENLPSPSICS 18
 | | | | | | | | | |
 Db 8 qckenlpsslcs 18

RESULT 3

AAB03185

ID AAB03185 standard; peptide; 103 AA.

AC AAB03185;

XX 23-OCT-2000 (first entry)

XX Human selenoprotein P C-terminal 103 residues.

XX Selenoprotein P fragment; human; programmed cell death inhibitor;
 KW apoptosis inhibitor; antiapoptotic; acquired immunodeficiency syndrome;
 KW AIDS; Parkinson's disease; Alzheimer's disease; myocardial infarct;
 KW cerebral infarct; stroke; reperfusion injury; organ transplant rejection;
 KW arteriosclerosis; redox disorder; immune disorder; cardiovascular;
 KW cerebroprotective; culture medium additive.

XX Homo sapiens.

Key	Location/Qualifiers
Modified-site 22	/note= "Selenocysteine"
Modified-site 40	/note= "Selenocysteine"
Modified-site 52	/note= "Selenocysteine"
Modified-site 67	/note= "Selenocysteine"
Modified-site 74	/note= "Selenocysteine"
Modified-site 89	/note= "Selenocysteine"
Modified-site 91	/note= "Selenocysteine"
Modified-site 98	/note= "Selenocysteine"
Modified-site 100	/note= "Selenocysteine"

Db 319 qckenlpslcs 329

RESULT 5
AAG75466

ID AAG75466 standard; Protein; 305 AA.
AC AAG75466;
DT 03-SEP-2001 (first entry)
DE Human colon cancer antigen protein SEQ ID NO:6230.
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
colorectal carcinoma.
XX Homo sapiens.
OS
PN WO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26524.
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI; 2001-235357/24.
DR N-PSDB; AAH34871.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
useful for preventing, diagnosing and/or treating colorectal cancers -
XX
PS Claim 11; Page 7679-7680; 9803pp; English.
XX
AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
cancer-associated nucleic acid molecules (N) and proteins (P), where
the proteins are collectively known as colon cancer antigens. The colon
cancer antigens have cytostatic activity and can be used in gene
therapy and vaccine production. N and P may be used in the prevention,
diagnosis and treatment of diseases associated with inappropriate P
expression. For example, N and P may be used to treat disorders
associated with decreased expression by rectifying mutations or deletions
in a patient's genome that affect the activity of P by expressing
inactive proteins or to supplement the patients own production of P.
Additionally, N may be used to produce the colon cancer-associated Ps,
by inserting the nucleic acids into a host cell and culturing the cell
to express the proteins. N and P can be used in the prevention, diagnosis
and treatment of colorectal carcinomas and cancers. AAH37195 to AAH37204
and AAG77789 represent sequences used in the exemplification of the
present invention.
N.B. Pages 666 to 682 and page 7053 of the sequence listing were
missing at time of publication, meaning no sequences are present for
SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 305 AA;

Query Match 25.0%; Score 7; DB 22; Length 305;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 QGLRAEE 26
|||||||
Db 32 qglraee 38

RESULT 6
ABG23820

ID ABG23820 standard; Protein; 832 AA.
XX
AC ABG23820;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #23811.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS88007.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity -
XX
PS Claim 20; SEQ ID No 54179; 103pp; English.
XX
The invention relates to isolated polynucleotide (I) and
polypeptide (II) sequences. (I) is useful as hybridisation probes,
polymerase chain reaction (PCR) primers, oligomers, and for chromosome
and gene mapping, and in recombinant production of (II). The
polynucleotides are also used in diagnostics as expressed sequence tags
for identifying expressed genes. (I) is useful in gene therapy techniques
to restore normal activity of (II) or to treat disease states involving
(II). (II) is useful for generating antibodies against it, detecting or
quantitating a polypeptide in tissue, as molecular weight markers and as
a food supplement. (II) and its binding partners are useful in medical
imaging of sites expressing (II). (I) and (II) are useful for treating
disorders involving aberrant protein expression or biological activity.
The polypeptide and polynucleotide sequences have applications in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits to assess biodiversity
and to produce other types of data and products dependent on DNA and
amino acid sequences. ABG00010-ABG30377 represent novel human
diagnostic amino acid sequences of the invention.
Note: The sequence data for this patent did not appear in the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 832 AA;

Query Match 25.0%; Score 7; DB 22; Length 832;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 CKENLPS 15
|||||||
Db 815 ckenlps 821

RESULT 7
ABG17156
ID ABG17156 standard; Protein; 932 AA.

XX AC ABG17156;
XX AC
XX DT 18-FEB-2002 (first entry)
XX DE
XX DE Novel human diagnostic protein #17147.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX PI WPI: 2001-639362/73.
XX DR N-PSDB; AAS81343.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -
XX PT
XX PT Claim 20; SEQ ID No 47515; 103pp; English.
XX PS
XX The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human
XX CC diagnostic amino acid sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ
XX Sequence 932 AA;
Query Match 25.0%; Score 7; DB 22; Length 932;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9 CKENLPS 15
Db 800 ckenlps 806
|||||
RESULT 8
AAU64969
ID AAU64969 standard; Protein; 62 AA.
XX

AC AAU64969;
XX 27-FEB-2002 (first entry)
XX DE
XX DE Propionibacterium acnes immunogenic protein #25865.
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX KW dermatological; osteopathic; neuroprotectant.
XX OS
XX OS Propionibacterium acnes.
XX PN WO200181581-A2.
XX PD
XX PD 01-NOV-2001.
XX PF 20-APR-2001; 2001WO-US12865.
XX PR 21-APR-2000; 2000US-199047P.
XX PR 02-JUN-2000; 2000US-208841P.
XX PR 07-JUL-2000; 2000US-216747P.
XX PA (CORI-) CORIXA CORP.
XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI: 2001-616774/71.
XX DR N-PSDB; AAS59655.
XX PT
XX PT Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris -
XX PT
XX Example 1; SEQ ID No 26164; 1069pp; English.
XX PS
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX CC polypeptides. The proteins and their associated DNA sequences are used in
XX CC the treatment, prevention and diagnosis of medical conditions caused by
XX CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX CC P. acnes is also involved in infections of bone, joints and the central
XX CC nervous system, however it is particularly involved in the inflammatory
XX CC lesions associated with acne vulgaris. A method for detecting the
XX CC presence or absence of P. acnes in a patient comprises contacting a
XX CC sample with a binding agent that binds to the proteins of the invention
XX CC and determining the amount of bound protein in the sample. The
XX CC polypeptides may be used as antigens in the production of antibodies
XX CC specific for P. acnes proteins. These antibodies can be used to
XX CC downregulate expression and activity of P. acnes polypeptides and
XX CC therefore treat P. acnes infections. The antibodies may also be used as
XX CC diagnostic agents for determining P. acnes presence, for example, by
XX CC enzyme linked immunosorbent assay (ELISA).
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ
XX Sequence 62 AA;
Query Match 21.4%; Score 6; DB 22; Length 62;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13 LPSLCS 18
Db 32 lpslcs 37
|||||
RESULT 9
AAU06318
ID AAU06318 standard; Protein; 62 AA.

```

XX AC AAM06318;
XX DT 05-OCT-2001 (first entry)
XX DE Human foetal protein, SEQ ID NO: 49.
XX KW Human; foetal protein; cytostatic; immunosuppressive; immunostimulant;
XX KW neotropic; neuroprotective; thrombolytic; osteopathic; antiinflammatory;
XX KW gene therapy; antisense therapy; cancer; immune disorder;
XX KW growth disorder; osteoporosis; thrombolytic disorder;
XX KW nervous system disorder; inflammation.
XX OS Homo sapiens.
XX PN WO200155339-A2.
XX PD 02-AUG-2001.
XX PF 25-JAN-2001; 2001WO-US02723.
XX PR 25-JAN-2000; 2000US-0491404.
XX PR 15-SEP-2000; 2000US-0663870.
XX PR 06-NOV-2000; 2000US-0707351.
XX PA (HYSE-) HYSEQ INC.
XX PI Yeung G, Ford JE, Boyle BJ, Arterburn MC, Drmanac RA, Tang YT;
XX PI Liu C, Asundi V, Zhou P, Werhman T;
XX DR WPI; 2001-463571/50.
XX DR N-PSDB; AAH93993.
XX PT Novel fetal proteins useful for the treatment and diagnosis of diseases
XX PT associated with dysfunction of the protein e.g. cancers, immune
XX PT disorders, growth disorders, thrombolytic disorders, nervous system
XX PT disorders and inflammation.
XX PS Claim 10; Page 200; 715pp; English.
XX CC The invention relates to novel foetal polypeptides encoded by
XX CC polynucleotides comprising one of 477 sequences fully defined in the
XX CC specification. The foetal polynucleotides and polypeptides are
XX CC useful in the treatment and diagnosis of diseases such as cancers,
XX CC immune disorders, growth disorders (e.g. osteoporosis), thrombolytic
XX CC disorders, nervous system disorders and inflammation. The present
XX CC sequence is a polypeptide encoded by a cDNA assembled using
XX CC an expressed sequence tag (EST) found to be expressed in human
XX CC foetal tissue cDNA libraries.
XX SQ Sequence 62 AA;

Query Match 21.4%; Score 6; DB 22; Length 62;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LPSLCS 18
Db 49 lpslcs 54
|||||

RESULT 10
AAU67435
ID AAU67435 standard; Protein; 97 AA.
XX AC AAU67435;
XX DT 27-FEB-2002 (first entry)
XX DE Propionibacterium acnes immunogenic protein #28331.
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;

KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX OS Propionibacterium acnes.
XX PN WO200181581-A2.
XX PD 01-NOV-2001.
XX PF 20-APR-2001; 2001WO-US12865.
XX PR 21-APR-2000; 2000US-199047P.
XX PR 02-JUN-2000; 2000US-208841P.
XX PR 07-JUL-2000; 2000US-216747P.
XX PA (CORI-) CORIXA CORP.
XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX DR WPI; 2001-616774/71.
XX DR N-PSDB; AAS59803.
XX PT Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris.
XX PS Example 1; SEQ ID NO 28630; 1069pp; English.
XX CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX CC polypeptides. The proteins and their associated DNA sequences are used in
XX CC the treatment, prevention and diagnosis of medical conditions caused by
XX CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX CC P. acnes is also involved in infections of bone, joints and the central
XX CC nervous system, however it is particularly involved in the inflammatory
XX CC lesions associated with acne vulgaris. A method for detecting the
XX CC presence or absence of P. acnes in a patient comprises contacting a
XX CC sample with a binding agent that binds to the proteins of the invention
XX CC and determining the amount of bound protein in the sample. The
XX CC polypeptides may be used as antigens in the production of antibodies
XX CC specific for P. acnes proteins. These antibodies can be used to
XX CC downregulate expression and activity of P. acnes polypeptides and
XX CC therefore treat P. acnes infections. The antibodies may also be used as
XX CC diagnostic agents for determining P. acnes presence, for example, by
XX CC enzyme linked immunosorbent assay (ELISA).
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 97 AA;

Query Match 21.4%; Score 6; DB 22; Length 97;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LPSLCS 18
Db 67 lpslcs 72
|||||

RESULT 11
AAU31563
ID AAU31563 standard; Protein; 107 AA.
XX AC AAU31563;
XX DT 18-DEC-2001 (first entry)
XX DE Novel human secreted protein #2054.
XX KW

```

KW Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200179449-A2.
 XX
 PD 25-OCT-2001.
 XX
 PF 16-APR-2001; 2001WO-US08656.
 XX
 PR 18-APR-2000; 2000US-0552929.
 PR 26-JAN-2001; 2001US-0770160.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-611725/70.
 XX
 PT Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy -
 XX
 PS Claim 20; Page 458; 765pp; English.
 CC
 CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.
 XX
 SQ Sequence 107 AA;

Query Match 21.4%; Score 6; DB 22; Length 107;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LPSLCS 18
 Db 47 lpslcs 52
 |||||

RESULT 12
 AAU52347
 ID AAU52347 standard; Protein; 120 AA.
 XX
 AC AAU52347;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #13243.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX

PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI; 2001-616774/71.
 DR N-PSDB; AAS59554.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 PS Example 1; SEQ ID NO 13542; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 120 AA;

Query Match 21.4%; Score 6; DB 22; Length 120;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GURAE 26
 Db 97 girae 102
 |||||

RESULT 13
 ABB62732
 ID ABB62732 standard; Protein; 166 AA.
 XX
 AC ABB62732;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 14988.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.

XX PD 27-SEP-2001. PR 01-APR-1999; 99US-0127462.
XX PF 23-MAR-2001; 2001WO-US09231. PR 06-APR-1999; 99US-0128234.
XX PR 16-APR-1999; 99US-0128714.
XX PR 19-APR-1999; 99US-0129845.
XX PR 21-APR-1999; 99US-0130077.
XX PR 23-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PA (PEKE) PE CORP NY. PR 23-APR-1999; 99US-0130891.
XX PR 28-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PI Venter JC, Adams M, Li PWD, Myers EW; PR 30-APR-1999; 99US-0132407.
XX XX WPI; 2001-656860/75. PR 04-MAY-1999; 99US-0132484.
XX DR N-PSDB; ABL06835. PR 05-MAY-1999; 99US-0132485.
XX XX PR 06-MAY-1999; 99US-0132486.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more PR 06-MAY-1999; 99US-0132487.
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell PR 07-MAY-1999; 99US-0132863.
XX PT interactions - PR 11-MAY-1999; 99US-0134256.
XX PS Disclosure; SEQ ID NO 14988; 21pp + Sequence Listing; English. PR 14-MAY-1999; 99US-0134218.
XX XX PR 14-MAY-1999; 99US-0134221.
XX XX PR 14-MAY-1999; 99US-0134221.
XX CC The invention relates to an isolated nucleic acid detection reagent PR 14-MAY-1999; 99US-0134370.
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is PR 18-MAY-1999; 99US-0134768.
XX CC useful in developmental biology and in elucidating cell signalling and PR 19-MAY-1999; 99US-0134941.
XX CC cell-cell interactions in higher eukaryotes for the development of PR 20-MAY-1999; 99US-0135124.
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention PR 21-MAY-1999; 99US-0135353.
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA PR 24-MAY-1999; 99US-0135629.
XX CC sequences (ABL01840-ABL16175) and the encoded proteins PR 25-MAY-1999; 99US-0136021.
XX CC (ABB57737-ABB72072). PR 27-MAY-1999; 99US-0136392.
XX CC The sequence data for this patent did not form part of the printed PR 28-MAY-1999; 99US-0136782.
XX CC specification, but was obtained in electronic format directly from WIPO PR 01-JUN-1999; 99US-0137222.
XX CC at ftp.wipo.int/pub/published_pct_sequences. PR 03-JUN-1999; 99US-0137528.
XX XX PR 04-JUN-1999; 99US-0137502.
XX SQ Sequence 166 AA; PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.

Query Match 21.4%; Score 6; DB 22; Length 166;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13 LPSLCS 18
Db 65 lpslcs 70
|||||
RESULT 14
AAG06557
ID AAG06557 standard; Protein; 197 AA.
XX AC AAG06557;
XX XX 17-OCT-2000 (first entry)
XX DT Arabidopsis thaliana protein fragment SEQ ID NO: 3372.
XX DE Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX EP1033405-A2.
XX PN 06-SEP-2000.
XX PD 25-FEB-2000; 2000EP-0301439.
XX PF 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 23-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.


```
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145216.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 04-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.

PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.4%; Score 6; DB 21; Length 197;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21 GIRAEE 26
Db 114 girae 119
|||||

RESULT 15
ARG06556
ID AAG06556 standard; Protein; 258 AA.
XX
AC AAG06556;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 3371.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
```

PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142970.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144652.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158222.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.

PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.4%; Score 6; DB 21; Length 258;

Best Local Similarity 100.0%; Pred. No. 1.2e+02; Mismatches 0; Indels 0; Gaps 0;

QY 21 GLRAEE 26
|||||
Db 175 glraee 180

RESULT 16
AAB74477
ID AAB74477 standard; Protein; 266 AA.

XX AC AAB74477;
XX DT 30-MAY-2001 (first entry)
XX DE Wheat cis-prenyltransferase SEQ ID NO: 20.

XX KW Cis-prenyltransferase; plant defence; wheat; soybean; grape; rice;
XX KW marigold; rubber tree.

XX OS Triticum aestivum.
XX PN WO200121650-A2.

XX PD 29-MAR-2001.

XX PF 21-SEP-2000; 2000WO-US25856.

XX PR 21-SEP-1999; 99US-0155046.

XX PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX PI Coldren C, Flint D, Hallahan DL, Wang H;

XX DR WPI; 2001-266063/27.

XX DR N-PSDB; AAF79863.

XX PT Polynucleotides encoding plant cis-prenyltransferase proteins useful
PT for creating transgenic plants in which cis-prenyltransferase protein
PT is present at higher or lower levels than normal -

XX PS Claim 4; Page 74; 82pp; English.

XX CC The present invention provides the protein and coding sequences of
CC cis-prenyltransferases from rice, soybean, marigold, rubber tree, African
CC daisy, grape and wheat. These are likely to be involved in plant defence
CC mechanisms and can be used to modulate cis-prenyltransferase production
CC in plants. The present sequence is a cis-prenyltransferase.

SQ Sequence 266 AA;

Query Match 21.4%; Score 6; DB 22; Length 266;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
|||||
Db 24 qglrae 29

RESULT 17

AAG06555

ID AAG06555 standard; Protein; 273 AA.

XX AC AAG06555;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 3370.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 18-MAY-1999; 99US-0134370.

PR 19-MAY-1999; 99US-0134768.

PR 20-MAY-1999; 99US-0134941.

PR 21-MAY-1999; 99US-0135124.

PR 24-MAY-1999; 99US-0135353.

PR 25-MAY-1999; 99US-0135629.

PR 27-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.4%; Score 6; DB 21; Length 273;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 21 GLRAEE 26
|||||
Db 190 glraee 195

RESULT 18
AAW01142
ID AAW01142 standard; Protein; 280 AA.
XX
AC AAW01142;
XX
DT 18-DEC-1996 (first entry)
XX
DE TrpA protein.
XX
KW trpB; tryptophan synthase; L-tryptophan; essential; food;
KW trpA; stockbreeding; coryneform bacteria; transform.
XX
OS Brevibacterium flavum.
XX
PN JP08070871-A.
PD 19-MAR-1996.
XX
PF 05-SEP-1994; 94JP-0211319.
XX
PR 05-SEP-1994; 94JP-0211319.
XX
PA (MITU) MITSUBISHI CHEM CORP.
XX
DR WPI; 1996-203151/21.
DR N-PSDB; AAT40802.
XX
PT DNA contg. a gene encoding tryptophan synthase - useful for
PT L-tryptophan production in Coryneform bacteria
XX
PS Claim 1; Page 4-8; 8pp; Japanese.
XX
CC The present sequence that of the TrpA protein, part of a tryptophan
CC synthase. Coryneform bacteria transformed with DNA (AAT40802) encoding
CC it and TrpB protein, express the tryptophan synthase which is useful for
CC the prodn. of L-tryptophan, an essential amino acid usefull in foods,
CC drugs and stockbreeding.
XX
SQ Sequence 280 AA;

Query Match 21.4%; Score 6; DB 17; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TGSALT 6
|||||
Db 234 tgsalt 239

RESULT 19
AAM40933
ID AAM40933 standard; Protein; 280 AA.
XX
AC AAM40933;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 5864.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX

OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
PA
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AAI60089.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Example 2; SEQ ID NO 5864; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 280 AA;

Query Match 21.4%; Score 6; DB 22; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 13 LPSLCS 18
|||||
Db 135 lpslcs 140

RESULT 20
AAG93105
ID AAG93105 standard; Protein; 280 AA.
XX
AC AAG93105;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum protein fragment SEQ ID NO: 6859.
XX
KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
XX

OS Corynebacterium glutamicum.
 PN EP1108790-A2.
 XX
 PD 20-JUN-2001.
 XX
 PF 18-DEC-2000; 2000EP-0127688.
 XX
 PP 16-DEC-1999; 99JP-0377484.
 PR 07-APR-2000; 2000JP-0159162.
 PR 03-AUG-2000; 2000JP-0280988.
 XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 PI
 DR WPI; 2001-376931/40.
 DR N-PSDB; AAH68324.
 XX
 XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile or pattern of a gene and identifying homologous gene
 PT
 PS Claim 17; SEQ ID NO: 6859; 246pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from Coryneform bacterium, and identifying a homologue of a gene derived
 CC from Coryneform bacterium. Coryneform bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described
 CC in the exemplification of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.
 XX
 SQ Sequence 280 AA;

Query Match 21.4%; Score 6; DB 22; Length 280;
 Best local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
 |||||
 Db 234 tgsalt 239

RESULT 21
 AAB79855
 ID AAB79855 standard; Protein; 280 AA.
 XX
 AC AAB79855;
 XX
 DT 30-APR-2001 (first entry)
 XX
 DE Corynebacterium glutamicum MP protein sequence SEQ ID NO:444.
 XX
 KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;
 KW fine chemical production; microorganism; organic acid; nucleoside;
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
 KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
 KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN WO200100843-A2.
 XX

PD 04-JAN-2001.
 XX
 PF 23-JUN-2000; 2000WO-IB00923.
 XX
 PP 25-JUN-1999; 99US-0141031.
 PR 01-JUL-1999; 99DE-1030476.
 PR 02-JUL-1999; 99US-0142101.
 PR 08-JUL-1999; 99DE-1031415.
 PR 08-JUL-1999; 99DE-1031418.
 PR 08-JUL-1999; 99DE-1031419.
 PR 08-JUL-1999; 99DE-1031420.
 PR 08-JUL-1999; 99DE-1031424.
 PR 08-JUL-1999; 99DE-1031428.
 PR 08-JUL-1999; 99DE-1031434.
 PR 08-JUL-1999; 99DE-1031435.
 PR 08-JUL-1999; 99DE-1031443.
 PR 08-JUL-1999; 99DE-1031453.
 PR 08-JUL-1999; 99DE-1031457.
 PR 08-JUL-1999; 99DE-1031465.
 PR 08-JUL-1999; 99DE-1031478.
 PR 08-JUL-1999; 99DE-1031510.
 PR 08-JUL-1999; 99DE-1031541.
 PR 08-JUL-1999; 99DE-1031573.
 PR 08-JUL-1999; 99DE-1031592.
 PR 08-JUL-1999; 99DE-1031632.
 PR 08-JUL-1999; 99DE-1031634.
 PR 08-JUL-1999; 99DE-1031636.
 PR 09-JUL-1999; 99DE-1032125.
 PR 09-JUL-1999; 99DE-1032136.
 PR 09-JUL-1999; 99DE-1032130.
 PR 09-JUL-1999; 99DE-1032186.
 PR 09-JUL-1999; 99DE-1032206.
 PR 09-JUL-1999; 99DE-1032227.
 PR 09-JUL-1999; 99DE-1032228.
 PR 09-JUL-1999; 99DE-1032229.
 PR 09-JUL-1999; 99DE-1032230.
 PR 14-JUL-1999; 99DE-1032922.
 PR 14-JUL-1999; 99DE-1032926.
 PR 14-JUL-1999; 99DE-1032928.
 PR 14-JUL-1999; 99DE-1033004.
 PR 14-JUL-1999; 99DE-1033005.
 PR 14-JUL-1999; 99DE-1033006.
 PR 12-AUG-1999; 99US-0148613.
 PR 27-AUG-1999; 99DE-1040764.
 PR 27-AUG-1999; 99DE-1040765.
 PR 27-AUG-1999; 99DE-1040766.
 PR 27-AUG-1999; 99DE-1040832.
 PR 31-AUG-1999; 99DE-1041378.
 PR 31-AUG-1999; 99DE-1041379.
 PR 31-AUG-1999; 99DE-1041380.
 PR 31-AUG-1999; 99DE-1041394.
 PR 31-AUG-1999; 99DE-1041396.
 PR 03-SEP-1999; 99DE-1042076.
 PR 03-SEP-1999; 99DE-1042077.
 PR 03-SEP-1999; 99DE-1042079.
 PR 03-SEP-1999; 99DE-1042086.
 PR 03-SEP-1999; 99DE-1042087.
 PR 03-SEP-1999; 99DE-1042088.
 PR 03-SEP-1999; 99DE-1042095.
 PR 03-SEP-1999; 99DE-1042124.
 PR 03-SEP-1999; 99DE-1042129.
 PR 09-MAR-2000; 2000US-0187970.
 XX
 XX (BADI) BASF AG.
 PA
 XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
 PI
 XX WPI; 2001-137957/14.
 DR N-PSDB; AAF71974.
 DR
 XX Nucleic acids from Corynebacterium glutamicum encoding metabolic
 PT pathway proteins, useful for producing fine chemicals in
 PT microorganisms, including organic acids, nonproteinogenic amino acids,

PT and purine and pyrimidine bases -
 XX
 PS Claim 20; Page 804-805; 1737pp; English.

XX
 CC AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic pathway (MP) proteins given in AAF79634 to AAB80211. The C. glutamicum MP nucleic acids are useful for the production of fine chemicals in microorganisms, including organic acids, nonproteinogenic amino acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated and unsaturated fatty acids, diols, carbohydrates, aromatic compounds, vitamins, cofactors, polyketides and enzymes.

XX SQ Sequence 280 AA;

Query Match 21.4%; Score 6; DB 22; Length 280;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
 |||||
 Db 234 tgsalt 239

RESULT 22

AAAP70511
 ID AAP70511 standard; protein; 281 AA.

XX
 AC AAP70511;

XX DT 11-JAN-1991 (first entry)

XX TrpA enzyme.

XX L-tryptophan; amino acid; trpA.

XX JP62244382-A.

XX PD 24-OCT-1987.

XX PF 16-APR-1986; 86JP-0087600.

XX PR 16-APR-1987; 87JP-0087600.

XX (AJIN) AJINOMOTO KK.

XX WPI; 1987-339035/48.

XX Deoxyribonucleic acid, used for L-tryptophan production -
 PT containing linking site to synthesize proteins, site coding
 PT leader peptide, and site for tryptophan operon synthesis.

PS Disclosure; Fig 7; 20pp; Japanese.

XX The protein is encoded by the trp operon and is one of 5 or 6
 CC enzymes necessary for tryptophan biosynthesis.

CC See also AAP70506, AAP70507, AAP70508, AAP70509, AAP70510, AAP70511.

XX SQ Sequence 281 AA;

Query Match 21.4%; Score 6; DB 8; Length 281;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
 |||||
 Db 232 tgsalt 237

RESULT 23

AAAY73850
 ID AAY73850 standard; Protein; 310 AA.

XX
 AC AAY73850;
 XX
 DT 14-MAR-2000 (first entry)

XX Human prostate tumor EST fragment derived protein #37.

XX Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;
 KW treatment.

XX Homo sapiens.

XX DE19820190-A1.

XX 04-NOV-1999.

XX PF 28-APR-1998; 98DE-1020190.

XX PR 28-APR-1998; 98DE-1020190.

XX (META-) METAGEN GES GENOMFORSCHUNG MBH.

XX PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;

XX WPI; 1999-621386/54.

XX N-PSDB; AAZ52870.

XX PT New human nucleic acid sequences from pancreatic tumors, and related

XX proteins -

XX PS Claim 23; Page 324; 502pp; German.

XX This invention describes novel polypeptides and their encoding nucleic
 CC acids derived from human pancreatic tumor tissue which have cytostatic
 CC activity. The sequences are also useful in producing pharmaceutical
 CC compositions for treatment of pancreatic tumors. AAY73814-Y74252
 CC represent protein fragments encoded by the human pancreatic tumor cDNA
 CC library derived expressed sequence tag (EST) sequences represented in
 CC AAZ52858-253014.

XX SQ Sequence 310 AA;

Query Match 21.4%; Score 6; DB 20; Length 310;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
 |||||
 Db 152 qglrae 157

RESULT 24

AAAB92503
 ID AAB92503 standard; Protein; 332 AA.

XX
 AC AAB92503;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:10617.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.

PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.4%; Score 6; DB 21; Length 399;

Best Local Similarity 100.0%; Pred.No.1.8e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
|||||
Db 225 tgsalt 230

RESULT 26

AA991465

ID AA991465 standard; Protein; 404 AA.

XX AA991465;

AC AA991465;

XX

DT 29-JUN-2000 (first entry)

DE Human secreted protein sequence encoded by gene 15 SEQ ID NO:138.

XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;

KW antiHIV; antiinflammatory; nootropic; neuroprotective; antiallergic;

KW osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma;

KW antipsoriatic; cardiant; gene therapy; cancer; neurological disorder;

KW immune disease; inflammation; blood disorder; tumour.

XX

OS Homo sapiens.

XX

PN WO200006698-A1.

XX

PD 10-FEB-2000.

XX

PF 29-JUL-1999; 99WO-US17130.

XX

PR 30-JUL-1998; 98US-0094657.

PR 05-AUG-1998; 98US-0095486.

PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.

PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.4%; Score 6; DB 21; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
Db 257 tgsalt 262

RESULT 28
AAP90475
ID AAP90475 standard; peptide; 442 AA.

XX
AC AAP90475;

XX
DT 24-OCT-1989 (first entry)

XX
DE N-terminal dimer of a Mullerian inhibiting substance-like polypeptide.
XX Mullerian inhibiting substance; bovine; cancer;
KW contraception; N-terminal dimers.
XX Bovine (bos taurus).

OS
XX WO8906695-A.
PN

XX
PD 27-JUL-1989.

XX
PF 25-JAN-1989; 89WO-US00239.

XX
PR 25-JAN-1988; 88US-0281301.

XX (BIOJ) BIOGEN. INC.
XX PA Cate RL, Pepinsky RB;
XX PI
XX
XX WPI; 1989-233849/32.
DR N-PSDB; AAN90395.
XX
XX Cleaved dimers of Mullerian inhibiting substance like polypeptide(s)
PT - are useful in treating cancers and for contraception.
XX
XX Claim 7; page 37; 64pp; English.
XX
XX Monomer of an N-terminal dimer of bovine Mullerian inhibiting substance,
CC is linked to another monomer by disulphide bonds.
CC It is made by cleaving between Arg-Ser in the sequence
CC aqsag. The dimers are used to treat cancer, esp. female genital tract
CC cancer, and for contraception. See also: AAN90395 and AAP90476-8.
XX
XX Sequence 442 AA;
SQ

Query Match 21.4%; Score 6; DB 10; Length 442;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 QGLRAE 25
Db 423 qglrae 428
|||||

RESULT 29
AAB66866
ID AAB66866 standard; Protein; 443 AA.
XX
AC AAB66866;
XX
DT 10-APR-2001 (first entry)
XX
DE Human peptidyl-prolyl isomerase-2.
XX
KW Human; cytostatic; immunosuppressive; antiinflammatory; cancer;
KW peptidyl-prolyl isomerase-2; HPIP-2; autoimmune disease;
KW inflammatory disease.
XX
OS Homo sapiens.
XX
XX US6171843-B1.
PN
XX 09-JAN-2001.
PD
XX
XX 01-JUN-1998; 98US-0088425.
PF
XX
XX 01-JUN-1998; 98US-0088425.
PR
XX
PA (INCY-) INCYTE PHARM INC.
XX
XX Bandman O, Lal P, Corley NC, Patterson C, Baughn MR;
PI
XX WPI; 2001-122332/13.
DR
DR N-PSDB; AAF32045.
XX
XX New nucleic acids encoding human peptidyl-prolyl isomerases, useful for
PT treating, diagnosing and preventing cancer, autoimmune disorders, and
PT inflammatory disorders -
PT
XX Claim 1; Fig 2; 34pp; English.
XX
XX The present sequence is human peptidyl-prolyl isomerase-2 (HPIP-2).
CC HPIP-2 is useful for diagnosing, treating or preventing disorders
CC associated with expression of human peptidyl-prolyl isomerase, such as
CC cancer, autoimmune diseases and inflammatory diseases. The coding
CC sequence for HPIP-2 was first identified as Incyte Clone 289973 from

CC a normalised brain cDNA library (BRAINON01).
XX
SQ Sequence 443 AA;
XX

Query Match 21.4%; Score 6; DB 22; Length 443;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 QGLRAE 25
Db 286 qglrae 291
|||||

RESULT 30
ABB63146
ID ABB63146 standard; Protein; 469 AA.
XX
AC ABB63146;
XX
XX 26-MAR-2002 (first entry)
DT
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 16230.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
XX WO200171042-A2.
PN
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US09231.
PF
XX
XX 23-MAR-2000; 2000US-191637P.
PR
XX 11-JUL-2000; 2000US-0614150.
PR
XX (PEKE) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR
DR N-PSDB; ABL07249.
XX
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
PT
XX Disclosure; SEQ ID NO 16230; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 469 AA;
XX

Query Match 21.4%; Score 6; DB 22; Length 469;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 QGLRAE 25
Db 146 qglrae 151
|||||

RESULT 31
AAAY79306
ID AAAY79306 standard; Protein; 480 AA.

XX AC AAAY79306;
XX DT 18-JUL-2000 (first entry)
XX DE Synecocystis sp. ycf 24 gene protein product.

XX KW Infection; therapy; ycf 24 gene.
XX OS Synecocystis sp.

XX PN WO200016758-A2.

XX PD 30-MAR-2000.

XX PF 22-SEP-1999; 99WO-GB03180.

XX PR 22-SEP-1998; 98GB-0020658.

XX PA (MEDI-) MEDICAL RES COUNCIL.

XX PI Wilson RJM, Mullineaux CW, Law AE;

XX DR WPI; 2000-283425/24.

XX DR N-PSDB; AA294423.

XX PT Treatment of humans or non-human animals for infection by an organism
PT comprising the ycf 24 gene, especially a malaria parasite, by
PT administering an inhibitor of ycf 24 gene product expression or
PT activity -

XX PS Disclosure; Page 36-38; 43pp; English.

XX CC The present sequence is that of the Synecocystis sp. strain PCC6803
CC ycf 24 gene protein product. ycf24 is an essential gene that appears
CC to play a role in replication. Disruption of all copies of the ycf
CC 24 gene was shown to be lethal. Disruption of some copies of the
CC gene gave a 'ragged' phenotype and appeared to cause a delay in, or
CC interfere in, the process of septation during cell division. The
CC invention provides an inhibitor of ycf 24 gene product expression
CC and/or activity, and a method of inhibiting the growth of an organism
CC by contacting it with the inhibitor. Also provided is a method of
CC identifying a compound that inhibits the growth of an organism,
CC involving (i) contacting a test compound with the ycf 24 gene
CC product, and (ii) determining whether the test compound inhibits
CC the activity of, or binds to, the product. Compounds inhibiting
CC ycf gene product expression and/or activity (such as antibodies,
CC antisense sequences, chemical inhibitors etc.) can be administered
CC in vivo or ex vivo to inhibit the growth of e.g. protozoa, algae,
CC bacteria (e.g. Escherichia coli) and especially malaria parasites
CC (claimed). For example, inhibitors may be used ex vivo to treat
CC land, water or food. Inducible or tissue-specific promoters can
CC also be used to provide control where ycf 24 gene product
CC expression is inhibited, e.g. to limit inhibition to the duration
CC of a treatment, or to limit inhibition to affected cell types. The
CC inhibitors can also be included in medicaments for treating
CC infections.

XX SQ Sequence 480 AA;

Query Match 21.4%; Score 6; DB 21; Length 480;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
|||||

DB 306 tgsalt 311

RESULT 32
AAB26594
ID AAB26594 standard; protein; 481 AA.

XX AC AAB26594;

XX DT 01-FEB-2001 (first entry)

XX DE Synecocystis sp phytochrome-related gene Cph3.

XX KW Phytochrome; fluorescent apophytochrome; bilin; Cph.

XX OS Synecocystis sp.

XX PN WO200056355-A1.

XX PD 28-SEP-2000.

XX PF 14-MAR-2000; 2000WO-US06607.

XX PR 19-MAR-1999; 99US-0272809.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Lagarias JC;

XX DR WPI; 2000-602195/57.

XX PT Fluorescent apophytochrome-bilin conjugates, e.g. Cph2 (derived from
PT Synecocystis species) and phycoerythrobilin conjugate, useful as
PT fluorescent markers for biological research -

XX PS Example 1; Page 46; 52pp; English.

XX CC The present invention is related to fluorescent apophytochrome-bilin
CC conjugates, known as phytofluors. An apoprotein known as Cph2 from
CC Synecocystis species is used as the apophytochrome and the bilin is
CC preferably phycoerythrobilin. The phytofluors are useful as fluorescent
CC markers for biological research. The phytofluors have a long wavelength
CC absorption maxima, a high molar absorption coefficient and the
CC recombinant apoproteins can spontaneously assemble with a variety of
CC bilin chromophore precursors. The present sequence is a phytochrome
CC related protein from Synecocystis sp.

XX SQ Sequence 481 AA;

Query Match 21.4%; Score 6; DB 21; Length 481;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
|||||

DB 89 tgsalt 94

RESULT 33
ABB65103
ID ABB65103 standard; Protein; 482 AA.

XX AC ABB65103;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 22101.

XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.

XX OS Drosophila melanogaster.

XX Plowman GD, Martinez R, Whyte D, Sudersanam S;
 XX WPI; 2001-032161/04.
 DR N-PSDB; AAF44706.
 XX
 XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and
 PT treating immune-related diseases and disorders, cardiovascular disease,
 PT neurodegenerative diseases and/or cancers -
 XX
 PS Claim 10; Fig 1; 310pp; English.
 XX
 CC The present sequence is a novel protein kinase. The novel protein kinases
 CC and the nucleic acids that encode them may be used in the treatment and
 CC diagnosis of diseases associated with inappropriate kinase expression
 CC such as immune-related diseases and disorders, cardiovascular disease,
 CC neurodegenerative diseases and/or cancers. The nucleic acids and
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays. The kinase polypeptides may be used as antigens in the production
 CC of antibodies of kinase expression and activity. Anti-kinase antibodies
 CC and kinase antagonists may also be used to down regulate kinase
 CC expression and activity. Diseases related to kinase expression and
 CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
 CC disorders, complications of organ transplantation, myocardial infarction,
 CC immune disorders, cardiomyopathies, strokes, renal failure,
 CC oxidative-stress related disorders, chronic inflammatory bowel disease,
 CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
 CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
 CC reproductive disorders.
 XX
 SQ Sequence 505 AA;

Query Match 21.4%; Score 6; DB 22; Length 505;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 NLPsLC 17
 Db 221 nlpslc 226
 |||||

RESULT 38
 AAW62558
 ID AAW62558 standard; peptide; 508 AA.
 XX
 AC AAW62558;
 XX
 DT 25-SEP-1998 (first entry)
 XX
 DE Fibroblast growth factor receptor protein kinase substrate 2 (FRS2).
 KW Adapter protein; regulation; growth factor stimulation;
 KW fibroblast growth factor receptor protein kinase substrate 2; FRS2;
 KW cellular differentiation; cellular proliferation;
 KW stimulated fibroblast growth factor receptor; Ras/MAPK cascade;
 KW Grb-2/Sos complex; diagnosis; prevention; treatment.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 11..152
 FT /note= "phosphotyrosine binding domain"
 XX
 XX WO9824902-A1.
 XX
 XX 11-JUN-1998.
 PD
 XX
 XX 01-DEC-1997; 97WO-US21851.
 XX
 XX 03-DEC-1996; 96US-0032093.
 XX
 XX (SUGE-) SUGEN.

PA (UYNV-) UNIV NEW YORK MEDICAL CENT.
 XX
 PI Kouhara H, Lax I, Schilling JW, Schlessinger J;
 PI Spivak-Kroizman T;
 XX
 XX WPI; 1998-333322/29.
 DR
 XX New human adaptor protein FRS2 - useful in, e.g. diagnosis,
 PT prevention or treatment of abnormal conditions associated with cell
 PT proliferation or cell differentiation
 XX
 PS Claim 3; Fig 1A; 53pp; English.
 XX
 CC The present sequence represents an adaptor protein linking protein
 CC kinases to activating partners in cells. The protein is designated
 CC fibroblast growth factor receptor protein kinase substrate 2 (FRS2).
 CC It regulates the growth factor stimulation of cellular differentiation
 CC and cellular proliferation by linking stimulated fibroblast growth
 CC factor receptor to the Ras/MAPK cascade via Grb-2/Sos complex. The
 CC FRS2 protein is useful to diagnose abnormal conditions associated with
 CC cell proliferation or cell differentiation in organisms caused by an
 CC aberration in a signal transduction pathway characterised by an
 CC interaction between a FRS2 polypeptide and a natural binding partner
 CC (especially in mammals). Such abnormal conditions can be associated
 CC with cell proliferation or cell differentiation. The protein is also
 CC useful to identify compounds that enhance or block FRS2 interactions
 CC with natural binding partners, to diagnose, prevent or treat abnormal
 CC conditions.
 XX
 SQ Sequence 508 AA;

Query Match 21.4%; Score 6; DB 19; Length 508;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 ENLPSL 16
 Db 350 enlpsl 355
 |||||

RESULT 39
 AAY79307
 ID AAY79307 standard; Protein; 508 AA.
 XX
 AC AAY79307;
 XX
 DT 18-JUL-2000 (first entry)
 XX
 DE Escherichia coli ycf 24 gene protein product.
 KW Infection; therapy; ycf 24 gene.
 XX
 OS Escherichia coli.
 XX
 PN WO200016758-A2.
 XX
 XX 30-MAR-2000.
 PD
 XX
 PF 22-SEP-1999; 99WO-GB03180.
 XX
 PR 22-SEP-1998; 98GB-0020658.
 XX
 XX (MEDI-) MEDICAL RES COUNCIL.
 XX
 PI Wilson RJM, Mullineaux CW, Law AE;
 XX
 DR WPI; 2000-283425/24.
 DR N-PSDB; AAZ94424.
 XX
 XX Treatment of humans or non-human animals for infection by an organism
 PT comprising the ycf 24 gene, especially a malaria parasite, by
 PT administering an inhibitor of ycf 24 gene product expression or

PT activity -
PS Disclosure; Page 39-41; 43pp; English.
XX
CC The present sequence is that of the Escherichia coli ycf 24 gene
CC protein product. ycf 24 is an essential gene that appears to play
CC a role in replication; disruption of ycf 24 was shown to be lethal.
CC The invention provides an inhibitor of ycf 24 gene product expression
CC and/or activity, and a method of inhibiting the growth of an organism
CC by contacting it with the inhibitor. Also provided is a method of
CC identifying a compound that inhibits the growth of an organism,
CC involving (i) contacting a test compound with the ycf 24 gene
CC product, and (ii) determining whether the test compound inhibits
CC the activity of, or binds to, the product. Compounds inhibiting
CC ycf gene product expression and/or activity (such as antibodies,
CC antisense sequences, chemical inhibitors etc.) can be administered
CC in vivo or ex vivo to inhibit the growth of e.g. protozoa, algae,
CC bacteria (e.g. Escherichia coli) and especially malaria parasites
CC (claimed). For example, inhibitors may be used ex vivo to treat
CC land, water or food. Inducible or tissue-specific promoters can
CC also be used to provide control where ycf 24 gene product
CC expression is inhibited, e.g. to limit inhibition to the duration
CC of a treatment, or to limit inhibition to affected cell types. The
CC inhibitors can also be included in medicaments for treating
CC infections.
XX
SQ Sequence 508 AA;

Query Match 21.4%; Score 6; DB 21; Length 508;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
Db 334 tgsalt 339
|||||

RESULT 40
AAU04693
ID AAU04693 standard; protein; 508 AA.
XX
AC AAU04693;
XX
DT 26-SEP-2001 (first entry)
XX
DE Human suc-1-associated neurotrophic factor target protein (SNT-1).
XX
KW Suc-1-associated neurotrophic factor target protein; SNT-1;
KW protein co-ordinate data; human; fibroblast growth factor receptor; FGFR;
KW nuclear magnetic resonance; NMR spectroscopy; rational drug design;
KW three-dimensional structure.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 19..24
FT /note= "Beta sheet region #1"
FT Region 35..40
FT /note= "Beta sheet region #2"
FT Region 45..49
FT /note= "Beta sheet region #3"
FT Region 52..57
FT /note= "Beta sheet region #4"
FT Region 63..68
FT /note= "Beta sheet region #5"
FT Region 71..76
FT /note= "Beta sheet region #6"
FT Region 85..90
FT /note= "Beta sheet region #7"
FT Region 99..107
FT /note= "Alpha helix"
FT Region 111..115

FT /note= "Beta sheet region #8"
XX WO200151521-A2.
PN
XX
PD 19-JUL-2001.
XX
PD 10-JAN-2001; 2001WO-US00821.
XX
PR 12-JAN-2000; 2000US-0175867.
XX
PR 09-JAN-2001; 2001US-0757415.
XX
PA (MOUN) MOUNT SINAI SCHOOL MEDICINE.
XX
PI Zhou M;
XX
XX WPI; 2001-425868/45.
DR
XX New isolated nucleic acid sequences encoding polypeptides useful in
PT structure based rational drug design -
PT
XX Claim 7; Page 228-230; 235pp; English.
PS
XX The sequence represents the amino acid sequence of human suc-1-associated
CC neurotrophic factor target protein (SNT-1). The invention provides
CC fragments of SNT and fibroblast growth factor receptor (FGFR) which can
CC form a binding complex that is amenable to structural determinations by
CC nuclear magnetic resonance (NMR) spectroscopy. The invention provides
CC methodology for related structure based rational drug design using the
CC three-dimensional data. The invention fulfills the need for the
CC determination of the three-dimensional structure of such complexes as
CC described and for procedures for related structural base drug design
CC predicated on such structural data.
XX
SQ Sequence 508 AA;

Query Match 21.4%; Score 6; DB 22; Length 508;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ENLPSL 16
Db 350 enlpsl 355
|||||

RESULT 41
AA99653
ID AA99653 standard; Protein; 516 AA.
XX
AC AA99653;
XX
DT 03-NOV-2000 (first entry)
XX
DE Human GTPase associated protein-4.
XX
KW Guanine nucleotide binding protein; GTP-binding protein; G-protein;
KW GTPase; GTPase associated protein; GTPAP; cell proliferation;
KW autoimmune; inflammatory; immune system disorder; cancer; AIDS;
KW acquired immune deficiency syndrome; asthma; atherosclerosis;
KW arthritis; systemic lupus erythematosus; psoriasis; human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 3
FT /note= "Potential phosphorylation site"
FT Modified-site 13
FT /note= "Potential phosphorylation site"
FT Modified-site 14
FT /note= "Potential phosphorylation site"
FT Modified-site 42
FT /note= "Potential phosphorylation site"
FT Modified-site 96

FT		/note= "Potential phosphorylation site"	
FT	Modified-site	237	
FT		/note= "Potential phosphorylation site"	
FT	Modified-site	270	
FT		/note= "Potential phosphorylation site"	
FT	Modified-site	316	
FT		/note= "Potential phosphorylation site"	
FT	Modified-site	347	
FT		/note= "Potential phosphorylation site"	
FT	Modified-site	360	
FT		/note= "Potential phosphorylation site"	
FT	Modified-site	371	
FT		/note= "Potential phosphorylation site"	
FT	Modified-site	395	
FT		/note= "Potential phosphorylation site"	
FT	Modified-site	430	
FT		/note= "Potential phosphorylation site"	
FT	Modified-site	433	
FT		/note= "Potential phosphorylation site"	
FT	Modified-site	500	
FT		/note= "Potential phosphorylation site"	
XX			
PN	W02000031263-A2.		
XX			
PD	02-JUN-2000.		
XX			
PF	23-NOV-1999;	99WO-US28013.	
XX			
PR	23-NOV-1998;	98US-0109592.	
PR	04-FEB-1999;	99US-0118610.	
PR	06-APR-1999;	99US-0127990.	
XX			
PA	(INCY-) INCYTE PHARM INC.		
XX			
PI	Hillman JL, Tang YT, Bandman O, Lal P, Yue H, Lu DAM, Baughn MR,		
PI	Yang J, Azimzai Y;		
XX			
DR	WPI; 2000-400073/34.		
N-	PSDB; AAA49174.		
XX			
PT	Human GTPase associated proteins, polynucleotides, and antibodies,		
PT	useful for diagnosing, preventing and treating various diseases such as		
PT	atherosclerosis, cancer, acquired immune deficiency syndrome (AIDS),		
PT	asthma, and autoimmune diseases -		
XX			
PS	Claim 1; Page 85-86; 144pp; English.		
XX			
CC	Human cDNA libraries from various tissues were screened for GTPase		
CC	associated proteins (GTPAP). The present sequence is human		
CC	GTPAP-4 protein. This sequence was derived from a cDNA library of the		
CC	prostate tumour tissue from a 65 year old male. This protein is		
CC	expressed in reproductive, haematopoietic/immune and		
CC	gastrointestinal tissue. The GTPAP proteins may be used to define		
CC	agonists and antagonists of GTPAP activity and to generate antibodies		
CC	to GTPAP. This means the GTPAP proteins may be useful for treatment or		
CC	prevention of diseases associated with GTPAP such as cell proliferation		
CC	disorders, autoimmune disorders, inflammatory disorders, immune system		
CC	disorders, cancer, AIDS, asthma, atherosclerosis, arthritis, systemic		
CC	lupus erythematosus and psoriasis.		
XX			
SQ	Sequence 516 AA;		
	Query Match	21.4%; Score 6; DB 21; Length 516;	
	Best Local Similarity	100.0%; Pred. No. 2.3e+02;	
	Matches 6; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	20 QGLRAE 25		
Db	332 qglrae 337		
RESULT	42		

CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention.

XX Sequence 521 AA;

Query Match 21.4%; Score 6; DB 22; Length 521;

Best Local Similarity 100.0%; Pred. No. 2.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ENLPSL 16

|||||

Db 363 enlpsl 368

RESULT 43

AA99373

ID AAY99373 standard; Protein; 544 AA.

XX AC AAY99373;

XX AC AAY99373;

XX AC AAY99373;

DT 08-AUG-2000 (first entry)

XX Human PRO1374 (UNQ711) amino acid sequence SEQ ID NO:118.

XX Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;
KW transmembrane; secretion; immunoadhesion; pharmaceutical; screening.

XX Homo sapiens.

XX WO200012708-A2.

XX 09-MAR-2000.

XX 01-SEP-1999; 99WO-US20111.

XX 01-SEP-1998; 98US-0098716.

XX 01-SEP-1998; 98US-0098749.

XX 01-SEP-1998; 98US-0098750.

XX 02-SEP-1998; 98US-0098803.

XX 02-SEP-1998; 98US-0098821.

XX 02-SEP-1998; 98US-0098843.

XX 02-SEP-1998; 98US-0098536.

XX 02-SEP-1998; 98US-0099596.

XX 02-SEP-1998; 98US-0099598.

XX 02-SEP-1998; 98US-0099602.

XX 02-SEP-1998; 98US-0099642.

XX 02-SEP-1998; 98US-0099741.

XX 02-SEP-1998; 98US-0099754.

XX 02-SEP-1998; 98US-0099792.

XX 02-SEP-1998; 98US-0099808.

XX 02-SEP-1998; 98US-0099812.

XX 02-SEP-1998; 98US-0099815.

XX 02-SEP-1998; 98US-0100385.

XX 02-SEP-1998; 98US-0100388.

XX 02-SEP-1998; 98US-0100390.

XX 02-SEP-1998; 98US-0100584.

XX 02-SEP-1998; 98US-0100627.

XX 02-SEP-1998; 98US-0100661.

XX 02-SEP-1998; 98US-0100662.

XX 02-SEP-1998; 98US-0100664.

XX 02-SEP-1998; 98US-0100665.

XX 02-SEP-1998; 98US-0100668.

PR 17-SEP-1998; 98US-0100684.
PR 17-SEP-1998; 98US-0100710.
PR 17-SEP-1998; 98US-0100711.
PR 17-SEP-1998; 98US-0100919.
PR 17-SEP-1998; 98US-0100930.
PR 18-SEP-1998; 98US-0100848.
PR 18-SEP-1998; 98US-0100849.
PR 18-SEP-1998; 98US-0101014.
PR 18-SEP-1998; 98US-0101068.
PR 18-SEP-1998; 98US-0101071.
PR 22-SEP-1998; 98US-0101279.
PR 23-SEP-1998; 98US-0101471.
PR 23-SEP-1998; 98US-0101472.
PR 23-SEP-1998; 98US-0101474.
PR 23-SEP-1998; 98US-0101475.
PR 23-SEP-1998; 98US-0101476.
PR 23-SEP-1998; 98US-0101477.
PR 23-SEP-1998; 98US-0101479.
PR 24-SEP-1998; 98US-0101738.
PR 24-SEP-1998; 98US-0101741.
PR 24-SEP-1998; 98US-0101743.
PR 24-SEP-1998; 98US-0101915.
PR 24-SEP-1998; 98US-0101916.
PR 29-SEP-1998; 98US-0102207.
PR 29-SEP-1998; 98US-0102240.
PR 29-SEP-1998; 98US-0102307.
PR 29-SEP-1998; 98US-0102330.
PR 29-SEP-1998; 98US-0102331.
PR 30-SEP-1998; 98US-0102484.
PR 30-SEP-1998; 98US-0102487.
PR 30-SEP-1998; 98US-0102571.
PR 30-SEP-1998; 98US-0102570.
PR 01-OCT-1998; 98US-0102684.
PR 01-OCT-1998; 98US-0102687.
PR 02-OCT-1998; 98US-0102965.
PR 06-OCT-1998; 98US-0103258.
PR 06-OCT-1998; 98US-0103449.
PR 07-OCT-1998; 98US-0103314.
PR 07-OCT-1998; 98US-0103315.
PR 07-OCT-1998; 98US-0103328.
PR 07-OCT-1998; 98US-0103395.
PR 07-OCT-1998; 98US-0103396.
PR 07-OCT-1998; 98US-0103401.
PR 08-OCT-1998; 98US-0103633.
PR 08-OCT-1998; 98US-0103678.
PR 08-OCT-1998; 98US-0103679.
PR 08-OCT-1998; 98US-0103711.
PR 14-OCT-1998; 98US-0104257.
PR 20-OCT-1998; 98US-0104987.
PR 20-OCT-1998; 98US-0105000.
PR 20-OCT-1998; 98US-0105002.
PR 21-OCT-1998; 98US-0105104.
PR 22-OCT-1998; 98US-0105169.
PR 22-OCT-1998; 98US-0105266.
PR 26-OCT-1998; 98US-0105693.
PR 26-OCT-1998; 98US-0105694.
PR 27-OCT-1998; 98US-0105807.
PR 27-OCT-1998; 98US-0105881.
PR 27-OCT-1998; 98US-0105882.
PR 27-OCT-1998; 98US-0105882.
PR 28-OCT-1998; 98US-0106023.
PR 28-OCT-1998; 98US-0106029.
PR 28-OCT-1998; 98US-0106030.
PR 28-OCT-1998; 98US-0106032.
PR 28-OCT-1998; 98US-0106033.
PR 28-OCT-1998; 98US-0106178.
PR 29-OCT-1998; 98US-0106248.
PR 29-OCT-1998; 98US-0106384.
PR 29-OCT-1998; 98US-0108500.
PR 30-OCT-1998; 98US-0106464.
PR 03-NOV-1998; 98US-0106856.
PR 03-NOV-1998; 98US-0106902.
PR 03-NOV-1998; 98US-0106905.

CC susceptibility to tumour development, particularly adrenal, lung, colon,
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders.

SQ Sequence 544 AA;

Query Match 21.4%; Score 6; DB 22; Length 544;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSAT 6

Db 163 tgsait 168

RESULT 45

AAB66122
ID AAB66122 standard; protein; 544 AA.

XX AC

XX AAB66122;

DT 02-APR-2001 (first entry)

XX DE

XX Protein of the invention #34.

XX KW

XX Secreted; transmembrane; gene therapy.

XX OS

XX Unidentified.

XX PN

XX WO200078961-A1.

XX PD

XX 28-DEC-2000.

XX PF

XX 18-FEB-2000; 2000WO-US04342.

XX PR

XX 23-JUN-1999; 99US-0141037.

XX PR

XX 20-JUL-1999; 99US-0144758.

XX PR

XX 26-JUL-1999; 99US-0145698.

XX PR

XX 01-SEP-1999; 99WO-US20111.

XX PR

XX 29-OCT-1999; 99US-0162506.

XX PR

XX 30-NOV-1999; 99WO-US28313.

XX PR

XX 02-DEC-1999; 99WO-US28551.

XX PR

XX 16-DEC-1999; 99WO-US30095.

XX PR

XX 05-JAN-2000; 2000WO-US00219.

XX PR

XX 06-JAN-2000; 2000WO-US00376.

XX PA

XX (GEPH) GENENTECH INC.

XX PI

XX Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;

XX PI

XX Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;

XX PI

XX Pan J, Paooni NF, Roy MA, Smith V, Stewart TA, Tumas D;

XX PI

XX Watanabe CK, Williams PM, Wood WI;

XX XX

XX WPI; 2001-071395/08.

XX DR

XX Secreted and transmembrane proteins and nucleic acids designated PRO,

XX PT

XX useful as hybridization probes, in chromosome and gene mapping and gene

XX PT

XX therapy -

XX PS

XX Claim 1; Fig 68; 787pp; English.

XX XX

XX The present invention relates to secreted and transmembrane proteins.

XX CC

XX These proteins and the DNA encoding them may be used as hybridization

XX CC

XX probes, in chromosome and gene mapping and in the generation of

XX CC

XX anti-sense RNA and DNA. They may also be used to generate either

XX CC

XX transgenic animals or knockout animals which are in turn useful for

XX CC

XX development and screening of therapeutically useful reagents.

XX CC

XX The nucleic acids may also be used in gene therapy.

XX XX

XX SQ

XX Sequence 544 AA;

Query Match 21.4%; Score 6; DB 22; Length 544;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSAT 6

Db 163 tgsait 168

RESULT 46

AAG63831
ID AAG63831 standard; Protein; 544 AA.

XX AC

XX AAG63831;

DT 26-NOV-2001 (first entry)

XX DE

XX Amino acid sequence of human alpha subunit of prolyl 4-hydroxylase.

XX DE

XX Human; alpha subunit; prolyl 4-hydroxylase; alpha (III) subunit;

XX KW

XX collagen; kidney fibrosis.

XX KW

XX OS

XX Homo sapiens.

XX XX

XX WO200168868-A2.

XX PN

XX 20-SEP-2001.

XX PD

XX PF

XX 15-MAR-2001; 2001WO-US08267.

XX PF

XX 15-MAR-2000; 2000US-0189373.

XX XX

XX (FIBR-) FIBROGEN INC.

XX PA

XX Kivirikko K, Myllyharju J, Kukkola L, Hietala R;

XX PI

XX WPI; 2001-570871/64.

XX DR

XX N-PSDB; AAH78017.

XX DR

XX New alpha subunit of prolyl 4-hydroxylase and polynucleotide encoding

XX PT

XX the subunit, useful for diagnosis, prevention and treatment of diseases

XX PT

XX and disorders associated with increased or decreased expression of the

XX PT

XX subunit -

XX XX

XX Claim 1; Fig 5; 75pp; English.

XX PS

XX The present sequence represents a human alpha subunit of prolyl

XX CC

XX 4-hydroxylase, designated alpha (III) subunit. The alpha (III)

XX CC

XX subunit is useful for the production of recombinant collagen, and

XX CC

XX for the diagnosis, prevention and treatment of various diseases and

XX CC

XX disorders associated with decreased or increased production of the

XX CC

XX subunit in specific tissues. The polynucleotide is a source of probes

XX CC

XX of diseases and disorders associated with increased or decreased

XX CC

XX expression or activity of various prolyl 4-hydroxylase enzymes and

XX CC

XX to identify alpha or beta prolyl 4-hydroxylase or fragments in

XX CC

XX tissue, e.g. biopsies from specific tissues, etc or other biological

XX CC

XX samples. Small molecules that modulate, regulate and inhibit prolyl

XX CC

XX 4-hydroxylase activity are useful for treating and preventing kidney

XX CC

XX fibrosis and various other diseases and disorders.

XX CC

XX SQ

XX Sequence 544 AA;

Query Match 21.4%; Score 6; DB 22; Length 544;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSAT 6

Db 163 tgsait 168

RESULT 47
AAU16301
ID AAU16301 standard; Protein; 549 AA.
XX
AC AAU16301;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human novel secreted protein, Seq ID 1254.
XX
KW Human; immunosuppressive; antiarthritic; antirheumatic;
KW cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
KW vulnary; secreted protein; rheumatoid arthritis;
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
KW cerebrovascular disorder; cerebral ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
KW corneal infection; wound healing; epithelial cell proliferation;
KW skin ageing; food additive; preservative; antiproliferative.
OS Homo sapiens.
XX
PN WO200155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01341.
XX
PR 31-JAN-2000; 2000US-0179055.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 06-SEP-2000; 2000US-0230439.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.

PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX
DR WPI; 2001-488783/53.
DR N-PSDB; AAS26288.
XX
XX New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -
XX
PS Claim 11; SEQ ID No 1254; 980pp; English.
XX
XX The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
CC are also used in diagnosing a pathological condition or susceptibility
CC to a pathological condition. Antibodies to the proteins can also
CC be used in alleviating symptoms associated with the disorders and in
CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
CC immunosorbent assays (ELISA). Disorders which are diagnosed or treated
CC include autoimmune diseases e.g. rheumatoid arthritis,
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
CC and ocular disorders e.g. corneal infection, and many other
CC disorders listed in the specification. The polypeptides can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention.

Query Match 21.4%; Score 6; DB 22; Length 549;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ENLPSL 16
Db 299 enlpsl 304
|||||

RESULT 48
AAR76503
ID AAR76503 standard; Protein; 551 AA.
XX
XX
AC AAR76503;
XX

DT 19-DEC-1995 (first entry)
XX
XX Cattle MIS mature protein.
XX
KW Mullerian inhibiting substance; MIS; ovary cancer; therapy.
XX
OS Bos taurus.
XX
PN US5427780-A.
XX
PD 27-JUN-1995.
XX
PF 30-OCT-1985; 85US-0792880.
XX
PR 28-OCT-1986; 86US-0923879.
PR 30-OCT-1985; 85US-0792880.
PR 25-APR-1991; 91US-0693764.
PR 05-OCT-1992; 92US-0957061.
XX
PA (BIOJ) BIOGEN INC.
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Cate RL, Donahoe PK;
XX
DR WPI; 1995-239996/31.
XX
XX Compsn. comprising human Mullerian Inhibiting substance - useful for
PT treating cancer, esp. ovarian cancer.
PT
XX
PS Claim 1; Columns 9-10; 41pp; English.
XX
XX Probes based on bovine MIS protein fragments were used to screen a
CC cattle testis cDNA library to isolate a clone (pX21) encoding cattle
CC MIS. The mature bovine MIS protein has the sequence given in
CC AAR76503.
XX
SQ Sequence 551 AA;
Query Match 21.4%; Score 6; DB 16; Length 551;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
Db 423 qglrae 428
|||||

RESULT 49
AAG30483
ID AAG30483 standard; Protein; 557 AA.
XX
XX
AC AAG30483;
XX
XX
DT 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 36452.
DE
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 24-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 22-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.

PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161932.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.4%; Score 6; DB 21; Length 557;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGSALT 6
Db 383 tgsalt 388
|||||

RESULT 50

AAP70195
ID AAP70195 standard; protein; 575 AA.

AC AAP70195;

XX 10-APR-1991 (first entry)

XX Sequence of bovine mullerian inhibiting substance (MIS)-like
DE polypeptide.

XX Anti-tumour; cancer therapy; ovarian cancer; contraceptive.

XX Bos taurus.

XX EP221761-A.

XX 13-MAY-1987.

XX 29-OCT-1986; 86EP-0308400.

XX 30-OCT-1985; 85US-0792880.

XX (BIOJ) BIOGEN NV.

PA (GEHO-) GEN HOSPITAL CORP.

XX Cate RL, Donahoe PK;

XX WPI; 1987-158390/23.

DR N-PSDB; AAN70306.

XX DNA sequence encoding Mullerian inhibiting substance-like

PT

PT polypeptide - for use in treatment of cancer, esp. cancer of the
PT female genital tract
XX Example; Fig 3; 39pp; English.
XX Bovine MIS protein from bovine testes was purified and the AA
CC sequence of several fragments detd. Based on these protein
CC sequences, several antisense oligonucleotide DNA probes were
CC synthesised. The probes were then used to screen a bovine cDNA
CC library. The insert of one bovine MIS cDNA clone (pS21) was used
CC to isolate the human MIS gene from a human cosmid library and
CC partial cDNA clone from a human cDNA library.

XX SQ Sequence 575 AA;

Query Match 21.4%; Score 6; DB 8; Length 575;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 QGLRAE 25

Db 447 qglrae 452
|||||

Search completed: August 22, 2002, 11:19:44
Job time: 171 sec

THIS PAGE BLANK (USPTO)

```

FT SE_CYS 375 375 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SE_CYS 377 377 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 83 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 365 365 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 370 370 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 380 AA; 42236 MW; 72F7031941F47212 CRC64;

Query Match 74.4%; Score 122; DB 1; Length 380;
Best Local Similarity 77.8%; Pred. No. 8.7e-10;
Matches 21; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 CINQLLCKLPTDSEAPRSXCCHRH 29
    ||||| : ||| |||||
Db 281 CINQLLCKLSEAPSSCCCHRH 307

RESULT 3
SELP_BOVIN
ID SELP_BOVIN STANDARD; PRT; 402 AA.
AC P49507; 019003;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Selenoprotein P-like protein precursor.
GN SEPP1 OR SELP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RX MEDLINE=95364621; PubMed=7637580;
RA Saijoh K., Saito N., Lee M.J., Fujii M., Kobayashi T., Sumino K.;
RT "Analysis of bovine selenoprotein p-like protein gene and availability
RT of metal responsive element (MRE) located in its promoter.";
RL Gene 199:211-217(1997).
CC -!- FUNCTION: IT CONSTITUTES A MAJOR SELENIUM POOL IN THE BRAIN AND
CC MAY PLAY AN IMPORTANT ROLE IN DEVELOPING AND/OR MODULATING THE
CC MORPHOLOGY OF NEURONS AND/OR GLIAL CELLS.
CC -!- SUBCELLULAR LOCATION: Extracellular (By similarity).
CC -!- TISSUE SPECIFICITY: BRAIN AND KIDNEY. MOST PROMINENTLY EXPRESSED
CC IN THE CEREBELLAR CORTEX, HIPPOCAMPUS AND OLFACTORY BULB.
CC -!- MISCELLANEOUS: THE SELENOCYSTEINES ARE ALL ENCODED BY THE OPAL
CC CODON, UGA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC -----
CC EMBL; D25220; BAA04949.2; -
CC DR EMBL; D88033; BAA23414.1; -
CC DR EMBL; D88031; BAA23414.1; JOINED.
CC DR EMBL; D88032; BAA23414.1; JOINED.
CC DR HSP; P04355; 4MT2
CC KW Glycoprotein; Signal; Selenium; Selenocysteine; Repeat.
CC SIGNAL 1 19 BY SIMILARITY.

```

```

FT CHAIN 20 402 SELENOPROTEIN P-LIKE PROTEIN.
FT SE_CYS 59 59
FT SE_CYS 297 297
FT SE_CYS 307 307
FT SE_CYS 338 338
FT SE_CYS 350 350
FT SE_CYS 363 363
FT SE_CYS 365 365
FT SE_CYS 372 372
FT SE_CYS 388 388
FT SE_CYS 390 390
FT SE_CYS 397 397
FT SE_CYS 399 399
FT DOMAIN 204 239 H-P REPEATS.
FT DOMAIN 260 266 POLY-HIS.
FT CONFLICT 178 181 SRPQ -> KALE (IN REF. 2).
FT CONFLICT 256 256 T -> P (IN REF. 2).
FT CONFLICT 282 282 L -> V (IN REF. 2).
FT CONFLICT 312 312 Y -> D (IN REF. 2).
SQ SEQUENCE 402 AA; 45018 MW; B7CF18751F80B8FF CRC64;

Query Match 70.7%; Score 116; DB 1; Length 402;
Best Local Similarity 72.4%; Pred. No. 6.1e-09;
Matches 21; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 KRCINOLLCKLPTDSEAPRSXCCHRH 29
    ||||| : ||| |||||
Db 299 KRCINOLLCKLPTDSEAPRSXCCHRH 327

RESULT 4
SELP_RAT
ID SELP_RAT STANDARD; PRT; 385 AA.
AC P25236;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Selenoprotein P precursor (Sep).
GN SEPP1 OR SELP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-41; 267-287 AND 316-327.
RC TISSUE=Liver;
RX MEDLINE=91244760; PubMed=2037562;
RA Hill K.E., Lloyd R.S., Yang J.-G., Read R., Burk R.F.;
RT "The cDNA for rat selenoprotein P contains 10 TGA codons in the open
RT reading frame.";
RL J. Biol. Chem. 266:10050-10053(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Kidney;
RX MEDLINE=95364621; PubMed=7637580;
RA Saijoh K., Saito N., Lee M.J., Fujii M., Kobayashi T., Sumino K.;
RT "Molecular cloning of cDNA encoding a bovine selenoprotein P-like
RT protein containing 12 selenocysteines and a (His-Pro) rich domain
RT insertion, and its regional expression.";
RL Brain Res. Mol. Brain Res. 30:301-311(1995).
RN [3]
RP REVIEW.
RX MEDLINE=95017128; PubMed=7931697;
RA Burk R.F., Hill K.E.;
RT "Selenoprotein P: A selenium-rich extracellular glycoprotein.";
RL J. Nutr. 124:1891-1897(1994).
CC -!- FUNCTION: MIGHT BE RESPONSIBLE FOR SOME OF THE EXTRACELLULAR
CC ANTIOXIDANT DEFENSE PROPERTIES OF SELENIUM OR MIGHT BE INVOLVED IN
CC THE TRANSPORT OF SELENIUM.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MADE IN THE LIVER AND SECRETED INTO THE
CC PLASMA.

```

```
CC -!- MISCELLANEOUS: THE SELENOCYSTEINES ARE ALL ENCODED BY THE OPAL
CC CODON, UGA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M63574; AAA42129.1; -
DR EMBL; D25221; BAA04950.2; -
DR PIR; A40380; OMRTSP.
KW Glycoprotein; Signal; Selenium; Selenocysteine; Plasma.
FT SIGNAL 1 19
FT CHAIN 20 385 SELENOPROTEIN P.
FT DOMAIN 244 252 POLY-HIS.
FT SE_CYS 59 59
FT SE_CYS 264 264
FT SE_CYS 282 282
FT SE_CYS 323 323
FT SE_CYS 335 335
FT SE_CYS 357 357
FT SE_CYS 371 371
FT SE_CYS 373 373
FT SE_CYS 380 380
FT SE_CYS 382 382
FT CARBOHYD 83 83
FT CARBOHYD 174 174
FT CARBOHYD 188 188
FT CARBOHYD 370 370
FT CARBOHYD 375 375
FT SEQUENCE 385 AA; 42614 MW; 013BDF4FB741E3E8 CRC64;
CC -----
Query Match 65.2%; Score 107; DB 1; Length 385;
Best Local Similarity 70.4%; Pred. No. 1e-07;
Matches 19; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 3 CINQLCKLPTDSELA PRSXCCCHRL 29
DB 286 CINQLCKLSESGAATSCCCHCHRL 312
-----
RESULT 5
Y734.CHLPN
ID Y734.CHLPN STANDARD; PRT; 324 AA.
AC Q927H1;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein CPN0734/CP0012/CPJ0734.
GN CPN0734 OR CP0012 OR CPJ0734.
OS Chlamydia pneumoniae (Chlamydia phila pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kallman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
```

```
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RL pneumoniae AR39.";
RN Nucleic Acids Res. 28:1397-1406(2000).
RP SEQUENCE FROM N.A.
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kohara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -!- SIMILARITY: BELONGS TO THE UPF0176 FAMILY.
CC -!- SIMILARITY: CONTAINS 1 RHODANESE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001655; AAD18873.1; -
DR EMBL; AE002165; AAF37908.1; -
DR TIGR; CP0012; -
DR InterPro; IPR001763; Rhodanese_domain.
DR Pfam; PF00581; Rhodanese; 1.
DR SMART; SM00450; RHOD; 1.
KW Hypothetical protein; Complete proteome.
FT DOMAIN 125 209 RHODANESE.
FT SEQUENCE 324 AA; 37445 MW; C541271AECE44B83 CRC64;
CC -----
Query Match 32.3%; Score 53; DB 1; Length 324;
Best Local Similarity 46.7%; Pred. No. 2.5;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
QY 13 TDSLAPRSXCCHCR 27
DB 236 SDPVPAPTAECCHCQ 250
-----
RESULT 6
LEP4_MYXXA
ID LEP4_MYXXA STANDARD; PRT; 335 AA.
AC O30387;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Type 4 prepilin-like proteins leader peptide processing enzyme
DE [Includes: Leader peptidase (EC 3.4.99.-) (prepilin peptidase); N-
DE methyltransferase (EC 2.1.1.-)].
GN PILD.
OS Myxococcus xanthus.
OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
OC Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.
OX NCBI_TaxID=34;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DK1622;
RA Wu S.S., Wu J., Cheng Y.L., Kaiser D.;
RT "The pilH gene encodes an ABC transporter required for type IV pilus
RL biogenesis in Myxococcus xanthus.";
CC -!- FUNCTION: CLEAVES TYPE-4 FIMBRIAL LEADER SEQUENCE AND METHYLATES
CC THE N-TERMINAL (GENERALLY PHE) RESIDUE (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A24; ALSO KNOWN AS TYPE
CC III LEADER PEPTIDASE FAMILY.
CC -----
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF003632; AAC36157.1; -.
 DR MEROPS; A24.001; -.
 DR InterPro; IPR000045; Peptidase_C20.
 DR Pfam; PF01478; Peptidase_C20; 1.
 DR PRINTS; PR00864; PREILNPTASE.
 KW Multifunctional enzyme; Hydrolase; Protease; Transferase;
 KW Methyltransferase; Transmembrane.
 FT TRANSMEM 13 33 POTENTIAL.
 FT TRANSMEM 105 125 POTENTIAL.
 FT TRANSMEM 131 151 POTENTIAL.
 FT TRANSMEM 206 226 POTENTIAL.
 FT TRANSMEM 258 278 POTENTIAL.
 FT TRANSMEM 299 319 POTENTIAL.
 SQ SEQUENCE 335 AA; 36767 MW; 27DF58A5FD2819FA CRC64;

Query Match 32.0%; Score 52.5; DB 1; Length 335;
 Best Local Similarity 33.3%; Pred. No. 3.1;
 Matches 9; Conservative 7; Mismatches 10; Indels 1; Gaps 1;

QY 4 INOLCKLPTDSELA-PRSKCCHRL 29
 DB 29 LNVVITARVPLDQSIIVNRSPRCRGHV 55

RESULT 7
 TPIS_RAT
 ID TPIS_RAT STANDARD; PRT; 248 AA.
 AC P48500;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Triosephosphate isomerase (EC 5.3.1.1) (TIM).
 GN TP11.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY;
 RA Yoon K.L., Guidotti G.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate = glycerone
 CC phosphate.
 CC -!- PATHWAY: PLAYS AN IMPORTANT ROLE IN SEVERAL METABOLIC PATHWAYS.
 CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; L36250; AAA42278.1; -.
 DR HSP; P00938; 1HT1.
 DR InterPro; IPR000652; Trioseph_isomerase.
 DR Pfam; PF00121; TIM; 1.
 DR ProDom; PD001005; Trioseph_isomrse; 1.
 DR PROSITE; PS00171; TIM; 1.
 KW Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;
 KW Pentose shunt.

FT INIT_MET 0 0 BY SIMILARITY.
 FT ACT_SITE 95 95 BY SIMILARITY.
 FT ACT_SITE 165 165 BY SIMILARITY.
 SQ SEQUENCE 248 AA; 26790 MW; AED3FB38B658AAB1 CRC64;

Query Match 31.4%; Score 51.5; DB 1; Length 248;
 Best Local Similarity 40.9%; Pred. No. 3.2;
 Matches 9; Conservative 6; Mismatches 2; Indels 5; Gaps 1;

QY 1 KRCINOLLC-----KLPTDSEL 17
 DB 18 KKCLGELICTLNAAKLPADTEV 39

RESULT 8
 VENV_DHV11
 ID VENV_DHV11 STANDARD; PRT; 521 AA.
 AC P27427;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Envelope glycoprotein precursor.
 GN P4.

OS Dhori virus (strain Indian/1313/61) (Dho).
 OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
 OC Thogoto-like viruses.
 OX NCBI_TaxID=11319;
 RN [1]

RP SEQUENCE FROM N.A.
 RA FREEDMAN-PAULSTICH E.Z., FULLER F.J.;
 RT "Nucleotide sequence of the tick-borne, orthomyxo-like
 RL Dhori/Indian/1313/61 virus envelope gene.";
 RL Virology 175:10-18(1990).

CC -!- FUNCTION: POSSIBLE ROLE IN ENDOCYTOTIC FUSION EVENTS DURING
 CC INFECTION.
 CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (POSSIBLY HOMODIMER).
 CC -!- SIMILARITY: TO THOGOTO VIRUS ENVELOPE GLYCOPROTEIN AND TO
 CC BACULOVIRUSES MAJOR ENVELOPE GLYCOPROTEIN (P64/P67).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M34002; AAA47907.1; -.
 DR PIR; A34679; VGIVDH.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 521 ENVELOPE GLYCOPROTEIN.
 FT TRANSMEM 501 517 POTENTIAL.
 FT CARBOHYD 44 44 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 396 396 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 521 AA; 58676 MW; 34582A42BF306C18 CRC64;

Query Match 31.4%; Score 51.5; DB 1; Length 521;
 Best Local Similarity 44.0%; Pred. No. 6.3;
 Matches 11; Conservative 3; Mismatches 8; Indels 3; Gaps 1;

QY 3 CINQLCKLPTDSEL---APRSXCC 24
 DB 88 CYNDLIPKSPTESELRWSKQCC 112

RESULT 9
 CNRD_CAEEL

ID LEP4_BACSU STANDARD; PRT; 248 AA.
 AC P15378;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Type 4 prepilin-like proteins leader peptide processing enzyme (late
 competence protein comC) [includes: leader peptidase (EC 3.4.99.-)
 (prepilin peptidase); N-methyltransferase (EC 2.1.1.-)]
 GN COMC.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90036690; PubMed=2553669;
 RA Mohan S., Aghlon J., Guillen N., Dubnau D.A.;
 RT "Molecular cloning and characterization of comC, a late competence
 gene of Bacillus subtilis.";
 RL J. Bacteriol. 171:6043-6051(1989).
 RN [2]
 RP SEQUENCE OF 1-46 AND 181-248 FROM N.A.
 RC STRAIN=168 / PY79;
 RX MEDLINE=93123172; PubMed=8419299;
 RA Margolis P.S., Driks A., Losick R.;
 RT "Sporeulation gene spoIIB from Bacillus subtilis.";
 RL J. Bacteriol. 175:528-540(1993).
 CC -!- FUNCTION: CLEAVES TYPE-4 FIMBRIAL LEADER SEQUENCE AND METHYLATES
 THE N-TERMINAL (GENERALLY PHE) RESIDUE (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -!- INDUCTION: BY SEVERAL PROTEINS INCLUDING SPOOA, SPOOH, SIN, ABRB,
 DEGS, DEGU, COMA, COMB AND COMK.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A24; ALSO KNOWN AS TYPE
 III LEADER PEPTIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M30805; AAA83365.1; -;
 DR EMBL; L04520; AAB59022.1; -;
 DR EMBL; L04519; AAB59025.1; -;
 DR EMBL; Z99118; CAB14767.1; -;
 DR PIR; A33490; A33490.
 DR MEROPS; A24.009; .
 DR Subtilist; BG10323; comC.
 DR InterPro; IPR000045; Peptidase_C20.
 DR Pfam; PF01478; Peptidase_C20; 1.
 KW Multifunctional enzyme; Hydrolase; Protease; Transferase;
 KW Methyltransferase; Transmembrane; Complete proteome.
 FT TRANSMEM 1 21 POTENTIAL.
 FT TRANSMEM 68 88 POTENTIAL.
 FT TRANSMEM 90 110 POTENTIAL.
 FT TRANSMEM 114 134 POTENTIAL.
 FT TRANSMEM 143 163 POTENTIAL.
 FT TRANSMEM 178 198 POTENTIAL.
 FT TRANSMEM 223 243 POTENTIAL.
 SQ SEQUENCE 248 AA; 26520 MW; 526653DCEDBD6097 CRC64;

Query Match 30.2%; Score 49.5; DB 1; Length 248;

Best Local Similarity 50.0%; Pred. No. 6.1; Mismatches 6; Indels 1; Gaps 1;

Matches 10; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

QY 9 CKLPTD-SELAPRSXCCHCR 27

I::I I::I I::I I::I

Db 21 CRIPLHSIIAPRSSCFPCR 40

RESULT 12
 PURL_MYCTU
 ID PURL_MYCTU STANDARD; PRT; 754 AA.
 AC P54876; O06631;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phosphoribosylformylglycinamide synthase II (EC 6.3.5.3) (FGAM
 synthase II).
 GN PURL OR RV0803 OR MT0823 OR MTCV07H7A.06C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=H37RV;
 RC MEDLINE=96425868; PubMed=8828210;
 RA Jackson M., Berthet F.-X., Otal I., Raugier J., Martin C.,
 Gicquel B., Guilhot C.;
 RT "The Mycobacterium tuberculosis purine biosynthetic pathway:
 isolation and characterization of the purC and purL genes.";
 RL Microbiology 142:2439-2447(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=H37RV;
 RC MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 Horsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: ATP + 5'-phosphoribosylformylglycinamide + L-
 glutamine + H(2)O = ADP + phosphate + 5'-
 phosphoribosylformylglycinamide + L-glutamate.
 CC -!- PATHWAY: DE NOVO PURINE BIOSYNTHESIS; FOURTH STEP.
 CC -!- SUBUNIT: HETERODIMER OF TWO SUBUNITS, PURQ AND PURL.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE FGAMS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U34956; AAB41455.1; -;
 DR EMBL; Z95618; CAB09108.1; -;
 DR EMBL; AE006972; AAK45065.1; -;
 DR TIGR; MT0823; -;
 DR Tuberculist; RV0803; -;
 DR InterPro; IPR000728; AIRS_related.
 DR Pfam; PF00586; AIRS; 2.
 DR Pfam; PF02769; AIRS_C; 2.

KW Purine biosynthesis; Ligase; ATP-binding; Complete proteome.

FT NP_BIND 116 127
FT CONFLICT 36 36 G -> D (IN REF. 1).
FT CONFLICT 61 61 K -> A (IN REF. 1).
FT CONFLICT 78 78 A -> V (IN REF. 1).
FT CONFLICT 81 81 A -> V (IN REF. 1).
FT CONFLICT 118 118 A -> P (IN REF. 1).
FT CONFLICT 464 464 E -> O (IN REF. 3).
SQ SEQUENCE 754 AA; 79507 MW; AFDB14694FDA5E6D CRC64;

Query Match 29.9%; Score 49; DB 1; Length 754;

Best Local Similarity 44.8%; Pred. No. 20;

Matches 13; Conservative 4; Mismatches 10; Indels 2; Gaps 1;

QY 2 RCINQLLCLPTDSELAPRSXCC--HCRH 28

DB 29 RRIRQILGRRTDTLAMYVMVSEHCY 57

RESULT 13

LEP4_AERHY

ID LEP4_AERHY STANDARD; PRT; 290 AA.

AC P45794;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Type 4 prepilin-like proteins leader peptide processing enzyme

DE [includes: Leader peptidase (EC 3.4.99.-) (Prepilin peptidase); N-

DE methyltransferase (EC 2.1.1.-).]

GN TAPD.

OS Aeromonas hydrophila.

OC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;

OC Aeromonas.

OX NCBI_TaxID=644;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-AH65;

RX MEDLINE=96417863; PubMed=8820654;

RA Pepe C.M., Eklund M.W., Strom M.S.;

RT "Cloning of an Aeromonas hydrophila type IV pilus biogenesis gene

cluster: complementation of plus assembly functions and

characterization of a type IV leader peptidase/N-methyltransferase

required for extracellular protein secretion.";

RL Mol. Microbiol. 19:857-869(1996).

CC -!- FUNCTION: CLEAVES TYPE-4 FIMBRIAL LEADER SEQUENCE AND METHYLATES

THE N-TERMINAL (GENERALLY PHE) RESIDUE. PROCESSES THE TAPA PILIN

PRECURSOR DURING MEMBRANE TRANSLOCATION.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane

(Probable).

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A24; ALSO KNOWN AS TYPE

III LEADER PEPTIDASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

DR EMBL; U20255; AAC43998.1; -.

DR MEROPS; A24.001; -.

DR InterPro; IPR000045; Peptidase_C20.

DR Pfam; PF01478; Peptidase_C20; 1.

DR PRINTS; PR00864; PREPILNPTASE.

KW Multifunctional enzyme; Hydrolase; Protease; Transferrase;

Methyltransferase; Transmembrane; Inner membrane.

FT TRANSMEM 14 34 POTENTIAL.

FT TRANSMEM 106 126 POTENTIAL.

FT TRANSMEM 130 150 POTENTIAL.

FT TRANSMEM 161 181 POTENTIAL.

FT TRANSMEM 185 205 POTENTIAL.

FT TRANSMEM 232 252 POTENTIAL.
FT TRANSMEM 261 281 POTENTIAL.
SQ SEQUENCE 290 AA; 32308 MW; EA584F277041A99B CRC64;

Query Match 29.3%; Score 48; DB 1; Length 290;

Best Local Similarity 47.1%; Pred. No. 11;

Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 12 PTDSELAPRSXCCHCRH 28

DB 63 PPYNLMVPRSCPHCNH 79

RESULT 14

IRR_HUMAN

ID IRR_HUMAN STANDARD; PRT; 1297 AA.

AC P14616; O60724;

DT 01-APR-1990 (Rel. 14, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Insulin receptor-related protein precursor (EC 2.7.1.112) (IRR)

DE (IR-related receptor).

GN INSR OR IRR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Kidney;

RX MEDLINE=99243372; PubMed=10226785;

RA Haenze J., Berthold A., Klammt J., Gallaher B., Siebler T.,

RA Kratzsch J., Elminger M., Kiess W.;

RT "Cloning and sequencing of the complete cDNA encoding the human

insulin receptor related receptor.";

RL Horm. Metab. Res. 31:77-79(1999).

RN [2]

RP SEQUENCE OF 30-1297 FROM N.A.

RX MEDLINE=89359245; PubMed=2768234;

RA Shier P., Watt V.M.;

RT "Primary structure of a putative receptor for a ligand of the insulin

family.";

RL J. Biol. Chem. 264:14605-14608(1989).

CC -!- FUNCTION: THIS RECEPTOR PROBABLY BINDS AN INSULIN RELATED PROTEIN

AND HAS A TYROSINE-PROTEIN KINASE ACTIVITY.

CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein

tyrosine phosphate.

CC -!- SUBUNIT: PROBABLE TETRAMER OF 2 ALPHA AND 2 BETA CHAINS LINKED BY

DISULFIDE BONDS. THE ALPHA CHAINS CONTRIBUTE TO THE FORMATION OF

THE LIGAND-BINDING DOMAIN, WHILE THE BETA CHAIN CARRY THE KINASE

DOMAIN.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-

CC PROTEIN KINASES.

CC -!- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

DR EMBL; AF064078; AAC17167.1; -.

DR EMBL; J05046; AAC31759.1; -.

DR PIR; B36502; B36502.

DR HSSP; P06213; IIRK.

DR MIM; 147671; -.

DR InterPro; IPR000049; EGFR_L.

DR InterPro; IPR000719; Euk_pkinase.

DR InterPro; IPR003961; FN_III.

Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 CINQLCKLPTDSELAPRS 21
I: | | | | | | | | |
Db 1062 CLKPLTKIPAGSLSPRS 1080

RESULT 18

TAT_SIVAG
ID ADO_BOVIN STANDARD: PRT: 1339 AA.
AC P48034;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Aldehyde oxidase (EC 1.2.3.1).
GN AOX1 OR AO.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 43-48; 187-218 AND 538-573.
RC TISSUE=Liver;
RX MEDLINE=96125080; PubMed=8537361;
RA Caizzi M.L., Raviolo C., Ghibaudi E., de Giola L., Salmona M.,
RA Cazaniga G., Kurosaki M., Terao M., Garattini E.;
RT "Purification, cDNA cloning, and tissue distribution of bovine liver
RT aldehyde oxidase.";
RL J. Biol. Chem. 270:31037-31045(1995).
CC -!- CATALYTIC ACTIVITY: An aldehyde + H(2)O + O(2) = an acid +
CC H(2)O(2).
CC -!- COPACITOR: FAD, MOLYBDOPTERIN, AND TWO 2FE-2S CLUSTERS.
CC -!- SUBUNIT: HOMODIMER.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN LIVER, LUNG AND
CC SPLEEN.
CC -!- PM: THE N-TERMINUS IS BLOCKED.
CC -!- SIMILARITY: TO 2FE-2S FERREDOXINS IN THE N-TERMINAL DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch)

EMBL; X87251; CAA60701.1; .
DR InterPro; IPR002888; 2Fe-2S_BD.
DR InterPro; IPR000564; 2Fe2S_ferredoxin.
DR InterPro; IPR000674; Aldxan_dh_C.
DR InterPro; IPR000572; Euk_oxidored_molyb.
DR InterPro; IPR001041; Ferredoxin.
DR InterPro; IPR002346; dehydrog_molyb.
DR Pfam; PF01315; Ald_xan_dh_C; 1.
DR Pfam; PF02738; Ald_xan_dh_C2; 1.
DR Pfam; PF00941; FAD_binding_5; 1.
DR Pfam; PF00111; fer2; 1.
DR Pfam; PF01799; fer2-2; 1.
DR PRODOM; PD186071; 2Fe-2S_BD; 1.
DR PROSITE; PS00197; 2FE2S_FERREDOXIN; 1.
DR PROSITE; PS00559; MOLYBDOPTERIN_EUK; 1.
KW Oxidoreductase; NAD; Molybdenum; Flavoprotein; FAD; Iron-sulfur.
FT METAL 44 44 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 49 49 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 52 52 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
SQ SEQUENCE 1339 AA; 147610 MW; 3CA7FF9D2806F655 CRC64;

Query Match 28.7%; Score 47; DB 1; Length 1339;

Best Local Similarity 44.4%; Pred. No. 62;

Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 3 CINQLCKLPTDSELA PR 20
I: | | | | | | | | |
Db 671 CVGQLVCAVIADSEVQAR 688

RESULT 19

TAT_SIVAG
ID TAT_SIVAG STANDARD: PRT: 119 AA.
AC P27982;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN TAT.
OS Simian immunodeficiency virus (AGM3 isolate) (SIV-AGM).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11730;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90232731; PubMed=2158689;
RA Balier M., Garber C., Mueller C., Cichutek K., Kurth R.;
RT "Complete nucleotide sequence of a simian immunodeficiency virus from
RT African green monkeys: a novel type of intragroup divergence.";
RL Virology 176:216-221(1990).
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC -!- MISCELLANEOUS: THIS IS AN AFRICAN GREEN MONKEY ISOLATE.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch)

EMBL; M30931; AAA91917.1; .
DR PIR; A46356; A46356.
DR InterPro; IPR001831; HIV_Tat.
DR Pfam; PF00539; Tat; 1.
KW Transcription regulation; Activator; RNA-binding; Nuclear protein;
KW AIDS.
SQ SEQUENCE 119 AA; 13390 MW; 58D52BB9FD0B648B CRC64;

Query Match 28.0%; Score 46; DB 1; Length 119;
Best Local Similarity 35.7%; Pred. No. 9.5;
Matches 10; Conservative 2; Mismatches 4; Indels 12; Gaps 2;

QY 1 KRCINQLCKLPTDSELA PRSXC-HCR 27
I: | | | | | | | | |
Db 24 KRCTNKCCKC-----CCYHCQ 40

RESULT 20

MLP2_DROME
ID MLP2_DROME STANDARD: PRT: 495 AA.
AC Q24400; Q9VI62;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Muscle LIM protein MLP84B.
GN LIM3 OR MLP84B OR CGI0699.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyroidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;

Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
[4]
RN SEQUENCE OF 19-33 AND 434-479, AND SUBUNITS.
RX MEDLINE-92338230; PubMed=1352993;
RA Fujii T., Kobayashi T., Honke K., Gasa S., Ishikawa M., Shimizu T.,
RA Makita A.;
RT "Proteolytic processing of human lysosomal arylsulfatase A.";
RL Biochim. Biophys. Acta 1122:93-98(1992).
[5]
RN PARTIAL SEQUENCE, AND 2-AMINO-3-OXOPROPIONIC ACID MODIFICATION SITE.
RX MEDLINE-95354208; PubMed=7628016;
RA Schmidt B., Selmer T., Ingendoh A., von Figura K.;
RT "A novel amino acid modification in sulfatases that is defective in
multiple sulfatase deficiency.";
RL Cell 82:271-278(1995).
[6]
RN X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE-98191325; PubMed=9521684;
RA Lukatela G., Krauss N., Theis K., Selmer T., Gieselmann V.,
RA von Figura K., Saenger W.;
RT "Crystal structure of human arylsulfatase A: the aldehyde function
and the metal ion at the active site suggest a novel mechanism for
sulfate ester hydrolysis.";
RL Biochemistry 37:3654-3664(1998).
[7]
RN REVIEW ON MLD VARIANTS.
RX MEDLINE-95170731; PubMed=7866401;
RA Gieselmann V., Zlotogora J., Harris A., Wenger D.A., Morris C.P.;
RT "Molecular genetics of metachromatic leukodystrophy.";
RL Hum. Mutat. 4:233-242(1994).
[8]
RN VARIANT MLD GLN-84.
RX MEDLINE-92344341; PubMed=1353340;
RA Kappler J., von Figura K., Gieselmann V.;
RT "Late-onset metachromatic leukodystrophy: molecular pathology in two
siblings.";
RL Ann. Neurol. 31:256-261(1992).
[9]
RN VARIANT MLD PHE-96.
RX MEDLINE-91328147; PubMed=1678251;
RA Gieselmann V., Fluharty A.L., Tonnesen T., von Figura K.;
RT "Mutations in the arylsulfatase A pseudodeficiency allele causing
metachromatic leukodystrophy.";
RL Am. J. Hum. Genet. 49:407-413(1991).
[10]
RN VARIANT MLD ASP-99.
RX MEDLINE-91206410; PubMed=1673291;
RA Kondo R., Wakamatsu N., Yoshino H., Fukuhara N., Miyatake T.,
RA Tsuji S.;
RT "Identification of a mutation in the arylsulfatase A gene of a
patient with adult-type metachromatic leukodystrophy.";
RL Am. J. Hum. Genet. 48:971-978(1991).
[11]
RN VARIANT MLD SER-122.
RX MEDLINE-94063853; PubMed=7902317;
RA Honke K., Kobayashi T., Fujii T., Gasa S., Xu M., Takamaru Y.,
RA Kondo R., Tsuji S., Makita A.;
RT "An adult-type metachromatic leukodystrophy caused by substitution of
serine for glycine-122 in arylsulfatase A.";
RL Hum. Genet. 92:451-456(1993).
[12]
RN VARIANT MLD SER-122.
RX MEDLINE-94063853; PubMed=7902317;
RA Honke K., Kobayashi T., Fujii T., Gasa S., Xu M., Takamaru Y.,
RA Kondo R., Tsuji S., Makita A.;
RT "An adult-type metachromatic leukodystrophy caused by substitution of
serine for glycine-122 in arylsulfatase A.";
RL Hum. Genet. 92:451-456(1993).
[13]
RN VARIANT MLD LEU-136.
RX MEDLINE-95163939; PubMed=7860068;
RA Kafert S., Heinisch U., Zlotogora J., Zlotogora J., Gieselmann V.;
RT "A missense mutation p136L in the arylsulfatase A gene causes
instability and loss of activity of the mutant enzyme.";

Hum. Genet. 95:201-204(1995).
[14]
RN VARIANT MLD ARG-245.
RX MEDLINE-93319632; PubMed=8101083;
RA Hasegawa Y., Kawame H., Eto Y.;
RT "Mutations in the arylsulfatase A gene of Japanese patients with
metachromatic leukodystrophy.";
RL DNA Cell Biol. 12:493-498(1993).
[15]
RN VARIANT MLD MET-274.
RX MEDLINE-94004907; PubMed=8104633;
RA Harvey J.S., Nelson P.V., Carey W.F., Robertson E.F., Morris C.P.;
RT "An arylsulfatase A (ARSA) missense mutation (T274M) causing late-
infantile metachromatic leukodystrophy.";
RL Hum. Mutat. 2:261-267(1993).
[16]
RN VARIANT MLD SER-309.
RX MEDLINE-93318834; PubMed=8101038;
RA Kreysing J., Bohne W., Bosenberg C., Marchesini S., Turpin J.C.,
RA Baumann N., von Figura K., Gieselmann V.;
RT "High residual arylsulfatase A (ARSA) activity in a patient with
late-infantile metachromatic leukodystrophy.";
RL Am. J. Hum. Genet. 53:339-346(1993).
[17]
RN VARIANT SER-350.
RX MEDLINE-90083282; PubMed=2574462;
RA Gieselmann V., Polten A., Kreysing J., von Figura K.;
RT "Arylsulfatase A pseudodeficiency: loss of a polyadenylation signal
and N-glycosylation site.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:9436-9440(1989).
[18]
RN VARIANT MLD LEU-426.
RX MEDLINE-93202658; PubMed=8095918;
RA Barth M.L., Fensom A., Harris A.;
RT "Prevalence of common mutations in the arylsulfatase A gene in
metachromatic leukodystrophy patients diagnosed in Britain.";
RL Hum. Genet. 91:73-77(1993).
[19]
RN VARIANTS MLD VAL-212; VAL-224 AND TYR-295.
RX MEDLINE-94154687; PubMed=7906588;
RA Barth M.L., Fensom A., Harris A.;
RT "Missense mutations in the arylsulfatase A genes of metachromatic
leukodystrophy patients.";
RL Hum. Mol. Genet. 2:2117-2121(1993).
[20]
RN VARIANTS MLD LEU-82; TYR-172; CYS-201; GLN-311; VAL-335 AND TRP-390.
RX MEDLINE-96047150; PubMed=7581401;
RA Barth M.L., Fensom A., Harris A.;
RT "Identification of seven novel mutations associated with
metachromatic leukodystrophy.";
RL Hum. Mutat. 6:170-176(1995).
[21]
RN CHARACTERIZATION OF VARIANTS MET-274 AND VAL-335.
RX MEDLINE-96303701; PubMed=8723680;
RA Hess B., Kafert S., Heinisch U., Wenger D.A., Zlotogora J.,
RA Gieselmann V.;
RT "Characterization of two arylsulfatase A missense mutations D335V and
T274M causing late infantile metachromatic leukodystrophy.";
RL Hum. Mutat. 7:311-317(1996).
[22]
RN VARIANTS MLD
RX MEDLINE-97245886; PubMed=9090526;
RA Draghia R., Letourneur F., Dragan C., Manicom J., Blanchot C.,
RA Kahn A., Poenaru L., Caillaud C.;
RT "Metachromatic leukodystrophy: identification of the first deletion in
exon 1 and of nine novel point mutations in the arylsulfatase A
gene.";
RL Hum. Mutat. 9:234-242(1997).
[23]
RN VARIANT MLD 406-SER--THR-408 DEL.
RX MEDLINE-98141126; PubMed=9490297;
RA Regis S., Filocamo M., Stroppiano M., Corsolini F., Caroli F.,
RA Gatti R.;

RT "A 9-bp deletion (2320del9) on the background of the arylsulfatase A
RT pseudodeficiency allele in a metachromatic leukodystrophy patient and
RT in a patient with nonprogressive neurological symptoms.";
RL Hum. Genet. 102:50-53(1998).

RN [24]
RP VARIANTS MLD PRO-135 AND SER-179.
RX MEDLINE=98260871; PubMed=9600244;
RA Gomez-Lira M., Perusi C., Mottes M., Pignatti P.F., Manfredi M.,
RA Rizzuto N., Salvati A.;
RT "Molecular genetic characterization of two metachromatic
RT leukodystrophy patients who carry the r799G mutation and show
RT different phenotypes; description of a novel null-type mutation.";
RL Hum. Genet. 102:459-463(1998).

RN [25]
RP VARIANTS MLD GLN-390 AND TYR-397.
RX MEDLINE=98112481; PubMed=9452102;
RA Coulter-Mackie M.B., Gagnier L.;
RT "Two novel mutations in the arylsulfatase A gene associated with
RT juvenile (R390Q) and adult onset (H397Y) metachromatic
RT leukodystrophy.";
RL Hum. Mutat. Suppl. 1:S254-S256(1998).

RN [26]
RP VARIANT HIS-496.
RX MEDLINE=98415722; PubMed=9744473;
RA Ricketts M.H., Foretz R.D., Manowitz P.;
RT "The R496H mutation of arylsulfatase A does not cause metachromatic

Query Match 28.0%; Score 46; DB 1; Length 507;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 19 PRSXCCHC 26
II IIII
DB 495 PRPACCHC 502

RESULT 22

ID SVE_METH STANDARD; PRT; 553 AA.
AC 026157;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glutamyl-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase)
DE (GluRS).
GN GLTX OR MTH51.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanothermobacter.
OX NCBI_TaxID=145262;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DELTA H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jivani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RJ J. Bacteriol. 179:7135-7155(1997).

CC -I- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) = AMP +
CC diphosphate + L-glutamyl-tRNA(Glu).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic.

CC -I- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL; AE000797; AAB84558.1; -
DR HSP; P00962; IGTR.
DR InterPro; IPR000924; trna-synt_1c.
DR InterPro; IPR001412; trna-synt_1.
DR Pfam; PF00749; trna-synt_1c; 1.
DR PRINTS; PR00987; TRNASYNTHGLJ.
DR PROSITE; PS00178; AA-TRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 103 113 "HIGH" REGION.
SQ SEQUENCE 553 AA; 63090 MW; C6DB8AE92C2FEE6B CRC64;

Query Match 28.0%; Score 46; DB 1; Length 553;
Best Local Similarity 42.9%; Pred. No. 38;
Matches 9; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

OY 9 CKLPTDSELAPRSXCCHRL 29
I: | | | | | |
DB 193 CRPEEFRELKNGRGEACHCRSL 213

RESULT 23

EZHL_HUMAN STANDARD; PRT; 747 AA.
ID Q92800; Q43287; Q14459;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Enhancer of zeste homolog 1 (ENX-2).
GN EZH1 OR KIAA0388.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97079676; PubMed=8921387;
RA Abel K.J., Brody L.C., Valdes J.M., Erdos M.R., McKinley D.R.,
RA Castilla L.H., Merajver S.D., Couch F.J., Friedman L.S.,
RA Ostermeyer E.A., Lynch E.D., King M.-C., Welcsh P.L.,
RA Osborne-Lawrence S., Spillman M., Bowcock A.M., Collins F.S.,
RA Weber B.L.;
RT "Characterization of EZH1, a human homolog of Drosophila Enhancer of
RT zeste near BRCA1.";
RL Genomics 37:161-171(1996).

RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98146265; PubMed=9473645;
RA Ogawa M., Hiraoka Y., Taniguchi K., Aiso S.;
RT "Cloning and expression of a human/mouse Polycomb group gene,
RT ENX-2/Enx-2.";
RL Biochim. Biophys. Acta 1395:151-158(1998).

RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97349984; PubMed=9205841;
RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 4:141-150(1997).

RN [4]
RP SEQUENCE OF 434-538 FROM N.A.
RX MEDLINE=96039267; PubMed=7490091;
RA Rommens J.M., Durocher F., McArthur J., Tonin P., Leblanc J.F.,
RA Allen T., Sanson C., Ferri L., Narod S., Morgan K., Simard J.;
RT "Generation of a transcription map at the HSD17B locus centromeric to

```
RT BRCA1 at 17q21.1;
RL Genomics 28:530-542(1995).
CC -!- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF GENE TRANSCRIPTION
CC AND CHROMATIN STRUCTURE.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: CONTAINS 1 SET DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE EZ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U03015; AAC50778.1; -.
DR EMBL: AB004818; BAA25019.1; -.
DR EMBL: AB002386; BAA20842.1; -.
DR EMBL: L38934; AAB59574.1; -.
DR MIM: 601674; -.
DR InterPro: IPR001719; AP_endonuclease_2.
DR InterPro: IPR001214; SET.
DR Pfam: PF00856; SET; 1.
DR SMART: SM00395; SANT; 1.
DR SMART: SM00317; SET; 1.
DR PROSITE: PS0280; SET; 1.
KW Transcription regulation; Nuclear protein; DNA-binding.
FT DOMAIN 491 496 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 524 606 CYS-RICH.
FT DOMAIN 619 730 SET.
FT CONFLICT 353 353 P -> S (IN REF. 2).
FT CONFLICT 488 488 N -> Y (IN REF. 1).
FT CONFLICT 532 535 DSTC -> EAL (IN REF. 2).
FT CONFLICT 591 602 ASEHWDCKVSC -> POSTGARFPV (IN REF. 2).
FT CONFLICT 700 747 VVMVNGDHRIGIFAKRAIQAGELEFFDYRSQADALKYGI
FT ERETDLV -> GESQ (IN REF. 2).
SQ SEQUENCE 747 AA; 85270 MW; 7CF52269CDA011B CRC64;

Query Match 28.0%; Score 46; DB 1; Length 747;
Best Local Similarity 35.5%; Pred. No. 50;
Matches 11; Conservative 3; Mismatches 9; Indels 8; Gaps 1;

QY 7 LLCKLPTDSELAP-----RSXCCHCRHL 29
DQ 477 LLKLTDELMPNQKKKKRHLWAAHCRKI 507
DQ 477 LLKLTDELMPNQKKKKRHLWAAHCRKI 507

RESULT 24
EZHL_MOUSE
ID EZHL_MOUSE STANDARD; PRT; 747 AA.
AC P70351; O9R089;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Enhancer of zeste homolog 1 (ENX-2).
GN EZH1 OR ENX2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97357309; PubMed-9214638;
RA Laible G., Wolf A., Dorn R., Reuter G., Nislow C., Lebersorger A.,
RA Popkin D., Pillus L., Jenuwein T.;
RT "Mammalian homologues of the Polycomb-group gene Enhancer of zeste
RT mediate gene silencing in Drosophila heterochromatin and at S.
RT cerevisiae telomeres.";
RL EMBO J. 16:3219-3232(1997).
RN [2]
```

```
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX MEDLINE-98146265; PubMed-9473645;
RA Ogawa M., Hiroaka Y., Taniguchi K., Also S.;
RT "Cloning and expression of a human/mouse polycomb group gene, ENX-
RT 2/Enx-2.";
RL Biochim. Biophys. Acta 1395:151-158(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-129/SV;
RX MEDLINE-99160476; PubMed-10051331;
RA Laible G., Haynes A.R., Lebersorger A., O'Carroll D., Mattei M.-G.,
RA Denny P., Brown S.D.M., Jenuwein T.;
RT "The murine polycomb-group genes eh1 and eh2 map close to hox gene
RT clusters on mouse chromosomes 11 and 6.";
RL Mamm. Genome 10:311-314(1999).
CC -!- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF GENE TRANSCRIPTION
CC AND CHROMATIN STRUCTURE.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN KIDNEY, ADRENAL
CC GLAND, TESTIS AND BRAIN.
CC -!- SIMILARITY: CONTAINS 1 SET DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE EZ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U60453; AAC53279.1; -.
DR EMBL: AB004817; BAA25018.1; ALT_INIT.
DR EMBL: AF104360; AAD54021.1; -.
DR MGD: MGI:1097895; Ezhl.
DR InterPro: IPR001719; AP_endonuclease_2.
DR InterPro: IPR001005; Myb_DNA_bind.
DR InterPro: IPR001214; SET.
DR Pfam: PF00856; SET; 1.
DR SMART: SM00395; SANT; 1.
DR SMART: SM00317; SET; 1.
DR PROSITE: PS0280; SET; 1.
KW Transcription regulation; Nuclear protein; DNA-binding.
FT DOMAIN 491 496 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 524 606 CYS-RICH.
FT DOMAIN 619 730 SET.
FT CONFLICT 446 446 H -> Y (IN REF. 2).
FT CONFLICT 452 452 N -> Y (IN REF. 2).
SQ SEQUENCE 747 AA; 85187 MW; 43CECFBDF3E49192 CRC64;

Query Match 28.0%; Score 46; DB 1; Length 747;
Best Local Similarity 35.5%; Pred. No. 50;
Matches 11; Conservative 3; Mismatches 9; Indels 8; Gaps 1;

QY 7 LLCKLPTDSELAP-----RSXCCHCRHL 29
DQ 477 LLKLTDELMPNQKKKKRHLWAAHCRKI 507
DQ 477 LLKLTDELMPNQKKKKRHLWAAHCRKI 507

RESULT 25
AD08_HUMAN
ID AD08_HUMAN STANDARD; PRT; 824 AA.
AC P78325;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ADAM 8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE domain 8) (Cell surface antigen MS2) (CD156a antigen) (CD156).
GN ADAM8 OR MS2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```



```
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA MEDLINE=97271556; PubMed=9126482;
RX Yoshiyama K., Higuchi Y., Kataoka M., Matsuura K., Yamamoto S.;
RT "CD156 (human ADAM8): expression, primary amino acid sequence, and
RL gene location.";
RL Genomics 41:56-62(1997).
CC -!- FUNCTION: POSSIBLE INVOLVEMENT IN EXTRAVASATION OF LEUKOCYTES.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: EXPRESSED ON NEUTROPHILS AND MONOCYTES.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC -!- SIMILARITY: CONTAINS 1 DISINTEGRIN DOMAIN.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD156 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd156.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D26579; BAA05626.1; -.
DR HSSP: P18619; 1FVL.
DR MEROPS: M12.208; -.
DR MIM: 602267; -.
DR InterPro: IPR001762; Disintegrin.
DR InterPro: IPR000361; EGF-like.
DR InterPro: IPR002870; Pep_M12B_propep.
DR InterPro: IPR001590; Reprolysin.
DR InterPro: IPR000130; Zn_M12Bptdse.
DR Pfam: PF00200; disintegrin; 1.
DR Pfam: PF01562; Pep_M12B_propep; 1.
DR Pfam: PF01421; Reprolysin; 1.
DR ProDom: PD000664; Disintegrin; 1.
DR SMART: SM00050; DISIN; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS0215; ADAM_MEPRO; 1.
DR PROSITE: PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE: PS0214; DISINTEGRIN_2; 1.
DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
DR K01 Hydroxylase; Metalloprotease; Zinc; Signal; Glycoprotein;
KW Transmembrane; Antigen.
KW SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 824 ADAM 8.
FT DOMAIN 17 655 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 656 676 POTENTIAL.
FT DOMAIN 677 824 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 200 400 METALLOPROTEASE.
FT DOMAIN 408 494 DISINTEGRIN-LIKE.
FT METAL 334 334 ZINC (CATALYTIC) (PROBABLE).
FT ACT_SITE 335 335 BY SIMILARITY.
FT METAL 338 338 ZINC (CATALYTIC) (PROBABLE).
FT METAL 344 344 ZINC (CATALYTIC) (PROBABLE).
FT DISULFID 310 395 BY SIMILARITY.
FT CARBOHYD 67 67 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 91 91 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 436 436 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 824 AA; 88673 MW; 5DF8E05F30DF479E CRC64;
```

Query Match 28.0%; Score 46; DB 1; Length 824;
Best Local Similarity 30.8%; Pred. No. 55;
Matches 8; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

QY 2 RCINOLLCKLPTDSELAPRSXCCHCR 27

Db 433 RCNNTTCLAEAGACQAHCCTCCQCK 458

```
RESULT 26
LEP4_BURPS
ID LEP4_BURPS STANDARD; PRT; 275 AA.
AC Q92F70;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Type 4 prelin-like proteins leader peptide processing enzyme
DE [includes: leader peptidase (EC 3.4.99.-) (Preplin peptidase); N-
DE methyltransferase (EC 2.1.1.-)].
GN GSPO.
OS Burkholderia pseudomallei (Pseudomonas pseudomallei).
OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC Burkholderia
OX NCBI_TaxID=28450;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1026B;
RX MEDLINE=99350433; PubMed=10419967;
RA Deshazer D., Brett P.J., Burtnick M.N., Woods D.E.;
RT "Molecular characterization of genetic loci required for secretion of
RL exoproducts in Burkholderia pseudomallei.";
RL J. Bacteriol. 181:4661-4664(1999).
CC -!- FUNCTION: CLEAVES TYPE-4 FIMBRIAL LEADER SEQUENCE AND METHYLATES
CC THE N-TERMINAL (GENERALLY PHE) RESIDUE (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A24; ALSO KNOWN AS TYPE
CC III LEADER PEPTIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF110186; AAD05189.1; -.
DR MEROPS: A24.001; -.
DR InterPro: IPR000045; Peptidase_C20.
DR Pfam: PF01478; Peptidase_C20; 1.
DR PRINTS: PR00864; PREPILNPTASE.
KW Multifunctional enzyme; Hydrolase; Protease; Transferase;
KW Methyltransferase; Transmembrane; Inner membrane.
FT TRANSMEM 1 21 POTENTIAL.
FT TRANSMEM 113 133 POTENTIAL.
FT TRANSMEM 149 169 POTENTIAL.
FT TRANSMEM 173 193 POTENTIAL.
FT TRANSMEM 196 216 POTENTIAL.
FT TRANSMEM 219 239 POTENTIAL.
FT TRANSMEM 254 274 POTENTIAL.
SQ SEQUENCE 275 AA; 28989 MW; 8C372590EFF512DF CRC64;
```

Query Match 27.7%; Score 45.5; DB 1; Length 275;
Best Local Similarity 52.6%; Pred. No. 24;
Matches 10; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 11 LPTDSEL-APRSXCCHCR 28

Db 49 LPARYNLCVPRSAQPCGH 67

RESULT 27

LMB2_RAT

ID LMB2_RAT STANDARD; PRT; 1801 AA.

AC P15800;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)


```
Db 1040 RTCNL---LGTDPQRCPSDILCHC 1061
RESULT 28
EPPI_MOUSE
ID EPPI_MOUSE STANDARD; PRT; 134 AA.
AC Q9DA01;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Eppin precursor (Epididymal protease inhibitor) (Serine protease
inhibitor-like with Kunitz and WAP domains 1).
GN SPINLW1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC STRAIN-BALB/c; TISSUE=Testis, and Epididymis;
RA Shivashanmugam P., Hall S.H., Hamil K.G., French F.S., O'Rand M.G.,
RA Richardson R.T.;
RT "Characterization of monkey and mouse Eppin, a protease inhibitor from
epididymis and testis.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RC SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
RA Saito T., Okazaki Y., Gofjorori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN EPIDIDYMIS AND TESTIS.
CC -!- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 WAP-TYPE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF346413; AAK31335.1; -
CC EMBL; AK006296; BAB24514.1; -
CC InterPro; IPR002223; Kunitz_BPTI.
CC InterPro; IPR002221; WAP.
CC Pfam; PF00014; Kunitz_BPTI; 1.
CC Pfam; PF00095; Wap; 1.
CC PRINTS; PR00759; BASICPTASE.
CC SMART; SM00131; KU; 1.
CC SMART; SM00217; WAP; 1..
CC PROSITE; PS00317; 4.DISULFIDE_CORE; 1.
CC PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE; PS0279; BPTI_KUNITZ_2; 1.
KW Serine protease inhibitor; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 134 EPPIN.
FT DOMAIN 29 73 WAP.
FT DOMAIN 77 127 BPTI/KUNITZ INHIBITOR.
FT DISULFID 33 61 BY SIMILARITY.
FT DISULFID 40 65 BY SIMILARITY.
FT DISULFID 48 60 BY SIMILARITY.
FT DISULFID 54 69 BY SIMILARITY.
FT DISULFID 77 127 BY SIMILARITY.
FT DISULFID 86 110 BY SIMILARITY.
FT DISULFID 102 123 BY SIMILARITY.
SQ SEQUENCE 134 AA; 15470 MW; DFFEBG3D4D4C427F CRC64;

Query Match 27.4%; Score 45; DB 1; Length 134;
Best Local Similarity 52.9%; Pred. No. 15;
Matches 9; Conservative 3; Mismatches 3; Indels 2; Gaps 1;

QY 1 KRCIN--QLLCKLPDTS 15
|:|:| | | | | |
DB 67 KKCLNPQQDICSPLPKDS 83

RESULT 29
STEL DROME
ID STEL DROME STANDARD; PRT; 172 AA.
AC P15021;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Stellate protein.
GN STE.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON-R; TISSUE=Testis;
RX MEDLINE=90169476; PubMed=1689686;
RA Livak K.J.;
RT "Detailed structure of the Drosophila melanogaster stellate genes and
their transcripts.";
RL Genetics 124:303-316(1990).
CC -!- FUNCTION: RESPONSIBLE FOR THE APPEARANCE OF PROTEINACEOUS STAR-
CC SHAPED CRYSTALS IN THE PRIMARY SPERMATOCYTES OF D.MELANOGASTER
CC MALES LACKING A Y CHROMOSOME.
CC -!- MISCELLANEOUS: THERE ARE MULTIPLE COPIES OF THE STELLATE GENE IN
CC FRUIT FLY.
CC -!- SIMILARITY: BELONGS TO THE CASEIN KINASE 2 BETA CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X15899; CAA33906.1; -
CC PIR; S08120; S08120.
CC PIR; S24397; S24397.
CC FlyBase; FBgn0003523; Ste.
CC InterPro; IPR000704; CAS_kinase_II.
CC Pfam; PF01214; CK_II_beta; 1.
CC PRINTS; PR00472; CASNKINASEII.
CC PROSITE; PS01101; CK2_BETA; 1.
KW Testis; Multigene family.
SQ SEQUENCE 172 AA; 19507 MW; C86304F591E76F8A CRC64;
```

```
Query Match      27.4%; Score 45; DB 1; Length 172;
Best Local Similarity 70.0%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 5 NOLLCKLPTD 14
   |||:||||
DB 21 NOFLCRVPTD 30

RESULT 30
YA28_SCHPO      STANDARD;      PRT;      632 AA.
AC Q09699;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Hypothetical 71.9 kDa protein C2F7.08C in chromosome I.
GN SPAC2F7.08C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Gentles S., Church C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: SOME, TO YEAST SNF5.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z50142; CAA90495.1; -
KW Hypothetical protein.
SQ SEQUENCE 632 AA; 71915 MW; 7153CFA105690953 CRC64;

Query Match      27.4%; Score 45; DB 1; Length 632;
Best Local Similarity 27.0%; Pred. No. 59;
Matches 10; Conservative 3; Mismatches 12; Indels 12; Gaps 1;

OY 1 KRCINOLLCKLPTD-----SELAPRSXCH 25
   |||: | | | | |
DB 553 KRCVIETFAKFPDNPENVIVKPAALNPAERMTVRICCH 589

RESULT 31
AD08_MOUSE
ID AD08_MOUSE      STANDARD;      PRT;      826 AA.
AC Q05910;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ADAM 8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE domain 8) (Cell surface antigen MS2) (Macrophage cysteine-rich
DE glycoprotein) (CD156 antigen).
GN ADAM8 OR MS2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR;
RA Yamamoto S., Yoshiyama K., Setoguchi M., Matsuura K., Higuchi Y.,
RA Akizuki S.;
```

Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.

[2]
SEQUENCE FROM N.A.
STRAIN=BALB/C; TISSUE=Liver;
MEDLINE=97364747; PubMed=9218457;
Kataoka M., Yoshiyama K., Matsuura K., Hijiya N., Higuchi Y.,
Yamamoto S.;
"Structure of the murine CD156 gene, characterization of its
promoter, and chromosomal location.";
J. Biol. Chem. 272:18209-18215(1997).

[3]
PRELIMINARY SEQUENCE FROM N.A.
STRAIN=ICR;
MEDLINE=91197896; PubMed=1982220;
Yoshida S., Setoguchi M., Higuchi Y., Akizuki S., Yamamoto S.;
"Molecular cloning of cDNA encoding MS2 antigen, a novel cell surface
antigen strongly expressed in murine monocytic lineage.";
Int. Immunol. 2:585-591(1990).

CC -!- FUNCTION: POSSIBLE INVOLVEMENT IN EXTRAVASATION OF LEUKOCYTES.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: MACROPHAGES.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
CC -!- SIMILARITY: CONTAINS 1 DISINTEGRIN DOMAIN.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

CC -----
EMBL; D10911; BAA31771.1; -
EMBL; X13335; CAA31712.1; -
HSSP; P18619; 1FVL.
MEROPS; M12.208; -
MGD; MGI:107825; Adam8.
InterPro; IPR001762; Disintegrin.
InterPro; IPR000561; EGF-like.
InterPro; IPR002870; Pep_M12B_propep.
InterPro; IPR001590; Reprolysin.
InterPro; IPR000130; Zn_MTpeptidse.
Pfam; PF00200; disintegrin; 1.
Pfam; PF01562; Pep_M12B_propep; 1.
Pfam; PF01421; Reprolysin; 1.
ProDom; PD000664; Disintegrin; 1.
SMART; SM00050; DISIN; 1.
SMART; SM00181; EGF; 1.
PROSITE; PS50215; ADAM_MEPRO; 1.
PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
PROSITE; PS50214; DISINTEGRIN_2; 1.
PROSITE; PS00022; EGF_1; UNKNOWN1.
PROSITE; PS01186; EGF_2; UNKNOWN1.
PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Transmembrane; Glycoprotein; Antigen; Zinc; Hydrolase;
KW Metalloprotease; Signal.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 826 ADAM 8.
FT DOMAIN 17 658 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 659 683 POTENTIAL.
FT DOMAIN 684 826 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 196 395 METALLOPROTEASE.
FT DOMAIN 403 489 DISINTEGRIN-LIKE.
FT METAL 329 329 ZINC (CATALYTIC) (PROBABLE).
FT ACT_SITE 330 330 BY SIMILARITY.
FT METAL 333 333 ZINC (CATALYTIC) (PROBABLE).
FT METAL 339 339 ZINC (CATALYTIC) (PROBABLE).
FT DISULFID 305 390 BY SIMILARITY.
FT CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 431 431 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 614 614 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 826 AA; 90046 MW; 3142CC81DDBADF9 CRC64;


```
RESULT 34
WT2_COLLI STANDARD; PRT; 63 AA.
AC P15787;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Metallothionein-II (MT-2).
OS Columba livia (Domestic pigeon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Columbiformes; Columbidae; Columba.
OX NCBI_TaxID=8932;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=90167121; PubMed=2407296;
RA Lin L.-Y., Lin W.C., Huang P.C.;
RT "Pigeon metallothionein consists of two species.";
RL Biochim. Biophys. Acta 1037:248-255(1990).
CC !- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
CC !- RESIDUES THAT BIND VARIOUS HEAVY METALS.
CC !- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:
CC !- FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA
CC !- DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11
CC !- CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE
CC !- BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.
CC !- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.
DR PIR; S08191; S08191.
DR HSSP; P02795; 2MHU.
DR InterPro; IPR003019; Metallthion_vert.
DR Pfam; PF00131; metalthio; 1.
DR PRINTS; PR00860; MTVERTBRATE.
DR PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.
KW Metal-binding; Metal-thiolate cluster; Chelation.
FT DOMAIN 1 30 BETA.
FT METAL 31 63 ALPHA.
FT METAL 6 6 CLUSTER B.
FT METAL 8 8 CLUSTER B.
FT METAL 14 14 CLUSTER B.
FT METAL 16 16 CLUSTER B.
FT METAL 20 20 CLUSTER B.
FT METAL 22 22 CLUSTER B.
FT METAL 25 25 CLUSTER B.
FT METAL 27 27 CLUSTER B.
FT METAL 30 30 CLUSTER B.
FT METAL 34 34 CLUSTER A.
FT METAL 35 35 CLUSTER A.
FT METAL 37 37 CLUSTER A.
FT METAL 38 38 CLUSTER A.
FT METAL 42 42 CLUSTER A.
FT METAL 45 45 CLUSTER A.
FT METAL 49 49 CLUSTER A.
FT METAL 51 51 CLUSTER A.
FT METAL 59 59 CLUSTER A.
FT METAL 61 61 CLUSTER A.
FT METAL 62 62 CLUSTER A.
SQ SEQUENCE 63 AA; 6452 MW; A2B84227E94C365E CRC64;

Query Match 26.8%; Score 44; DB 1; Length 63;
Best Local Similarity 34.8%; Pred. No. 10;
Matches 8; Conservative 3; Mismatches 8; Indels 8; Gaps 1;

QY 3 CINOLCKLPDSELA PRSXCH 25
| :||| :| :|||
Db 45 CAKGCVCKEPSSS----KSCCH 63

RESULT 35
AN11_HUMAN STANDARD; PRT; 84 AA.
Q9NYG5; Q9P0R2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Anaphase promoting complex subunit 11 (Hepatocellular carcinoma
DE associated ring finger protein) (Protein HSPC214).
GN ANAPC11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Chan A.H., Lee S.M.Y., Wayne M.M.Y., Fung K.P., Lee C.Y., Tsui S.K.W.;
RT "Characterization of a human hepatocellular carcinoma associated ring
RT finger protein.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=20499367; PubMed=11042152;
RA Zhang Q.-H., Ye M., Wu X.-Y., Ren S.-X., Zhao M., Zhao C.-J., Fu G.,
RA Shen Y., Fan H.-Y., Lu G., Zhong M., Xu X.-R., Han Z.-G., Zhang J.-W.,
RA Tao J., Huang Q.-H., Zhou J., Hu G.-X., Gu J., Chen S.-J., Chen Z.;
RT "Cloning and functional analysis of cDNAs with open reading frames for
RT 300 previously undefined genes expressed in CD34+ hematopoietic
RT stem/progenitor cells.";
RL Genome Res. 10:1546-1560(2000).
CC !- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC !- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC !- FRAMESHIFT IN POSITION 67.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF247565; AAF65816.1; .
DR EMBL; AF151048; AAF36134.1; ALT_FRAME.
DR InterPro; IPR001841; Znf_ring.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
DR PROSITE; PS50089; ZF_RING_2; 1.
KW Zinc-finger.
FT ZN_FING 34 77 RING-TYPE.
FT SEQUENCE 84 AA; 9841 MW; EACBD5A54FDC11AE CRC64;

Query Match 26.8%; Score 44; DB 1; Length 84;
Best Local Similarity 38.1%; Pred. No. 13;
Matches 8; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 9 CKLPDSELA PRSXCHCRHL 29
||:| | | | |
Db 37 CKVPGDDCLVWGQCSCFCFH 57

RESULT 36
AN11_MOUSE STANDARD; PRT; 84 AA.
ID AN11_MOUSE
AC Q9CPX9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Anaphase promoting complex subunit 11.
GN ANAPC11.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

```

```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fletschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Plesio G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AK003612; BAB22890.1; -
CC EMBL; AK003244; BAB22663.1; -
CC InterPro; IPR001841; Znf_Ring.
CC SMART; SM00184; RING; 1.
CC PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
CC PROSITE; PS00089; ZF_RING_2; 1.
CC Zinc-finger. 34 77 RING-TYPE.
FT ZN_FING 84 AA; 9618 MW; EACBD5A194FC11AE CRC64;
SQ SEQUENCE 84 AA; 9618 MW; 2; Mismatches 11; Indels 0; Gaps 0;

Query Match 26.8%; Score 44; DB 1; Length 84;
Best Local Similarity 38.1%; Pred. No. 13;
Matches 8; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 9 CKLPTDSELA PRSXCHCRHL 29
DB 37 CKVPGDCCPLVWGQCSHCFFM 57

RESULT 37
COL_CANFA COL_CANFA STANDARD; PRT; 112 AA.
AC P19090;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Colipase precursor.
GN CLPS.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=91016846; PubMed=2216731;
RA Fukuoka S.-I., Taniguchi Y., Kitagawa Y., Scheele G.;

```

```

RT "Full length cDNA sequence encoding canine pancreatic colipase.";
RL Nucleic Acids Res. 18:5549-5549(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=93266588; PubMed=7684378;
RA Fukuoka S.-I., Zhang D.E., Taniguchi Y., Scheele G.A.;
RT "Structure of the canine pancreatic colipase gene includes two
RT protein-binding sites in the promoter region.";
RL J. Biol. Chem. 268:11312-11320(1993).
CC -!- FUNCTION: COLIPASE IS A COFACTOR OF PANCREATIC LIPASE. IT ALLOWS
CC THE LIPASE TO ANCHOR ITSELF TO THE LIPID-WATER INTERFACE. WITHOUT
CC COLIPASE THE ENZYME IS WASHED OFF BY BILE SALTS, WHICH HAVE AN
CC INHIBITORY EFFECT ON THE LIPASE.
CC -!- FUNCTION: ENTEROSTATIN HAS A BIOLOGICAL ACTIVITY AS A SATIETY
CC SIGNAL.
CC -!- SUBUNIT: FORM A 1:1 STOICHIOMETRIC COMPLEX WITH PANCREATIC LIPASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X53564; CAA37636.1; -
CC EMBL; M63427; AAA03513.1; -
CC PIR; S11436; S11436.
CC PIR; A46717; A46717.
CC HSP; P02703; LPCN.
CC InterPro; IPR001981; Colipase.
CC Pfam; PF01114; Colipase; 1.
CC Pfam; PF02740; Colipase_C; 1.
CC PRINTS; PR00128; COLIPASE.
CC SMART; SM00023; COLIPASE; 1.
CC PROSITE; PS00121; COLIPASE; 1.
CC Lipid degradation; Digestion; Pancreas; Signal.
FT SIGNAL 1 17
FT PROPEP 18 22 ENTEROSTATIN, ACTIVATION PEPTIDE
FT CHAIN 23 112 COLIPASE
FT DISULFID 34 45 BY SIMILARITY.
FT DISULFID 40 56 BY SIMILARITY.
FT DISULFID 44 78 BY SIMILARITY.
FT DISULFID 66 86 BY SIMILARITY.
FT DISULFID 80 104 BY SIMILARITY.
SQ SEQUENCE 112 AA; 12035 MW; 96E5B821BA8CA71 CRC64;

Query Match 26.8%; Score 44; DB 1; Length 112;
Best Local Similarity 34.8%; Pred. No. 17;
Matches 8; Conservative 2; Mismatches 3; Indels 10; Gaps 1;

QY 3 CINQLCKLPTDSELA PRSXCH 25
DB 34 CLNSVQCK-----SKCH 46

RESULT 38
YNB7_YEAST
ID YNB7_YEAST STANDARD; PRT; 112 AA.
AC P53977;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Hypothetical 12.5 kDa protein in HDAL-PUBI intergenic region.
GN YNL017C OR N2834.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]

```

RP SEQUENCE FROM N.A.
RA Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RL
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; 271293; CAA95879.1; -
DR EMBL; 271292; CAA95878.1; -
DR SGD; S0004962; YNL017C.
KW Hypothetical protein.
SQ SEQUENCE 112 AA; 12537 MW; EFC0192B08F1AC29 CRC64;

Query Match 26.8%; Score 44; DB 1; Length 112;
Best Local Similarity 34.5%; Pred. No. 17;
Matches 10; Conservative 4; Mismatches 9; Indels 6; Gaps 1;

QY 2 RCINQL-----LCKLPTDSELA PRSXCC 24
DB 46 RCINQLGQETIVICFSPSLHTIKPTSMC 74

RESULT 39
TAT_SIVS4
ID TAT_SIVS4 STANDARD; PRT; 117 AA.
AC P12513;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN TAT
OS Simian immunodeficiency virus (F236/smH4 isolate) (sooty mangabey).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11737;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89262053; PubMed=2786147;
RA Hirsch V.M., Olmstead R.A., Murphy-Corb M., Purcell R.H.,
RA Johnson P.R.;
RT "An African primate lentivirus (SIVsm) closely related to HIV-2.";
RL Nature 339:389-392(1989).
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC
CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X14307; -; NOT_ANNOTATED_CDS.
DR HIV; X14307; TATSSMMH4.
DR InterPro; IPR001831; HIV_Tat.
DR Pfam; PF00539; Tat; 1.
DR PRINTS; PR00055; HIVTATDOMAIN.
KW Transcription regulation; Activator; RNA-binding; Nuclear protein;
KW AIDS.
SQ SEQUENCE 117 AA; 13405 MW; B5F71910631A8E45 CRC64;

Query Match 26.8%; Score 44; DB 1; Length 117;
Best Local Similarity 34.5%; Pred. No. 17;
Matches 10; Conservative 4; Mismatches 9; Indels 6; Gaps 1;

Best Local Similarity 33.3%; Pred. No. 18;
Matches 9; Conservative 2; Mismatches 4; Indels 12; Gaps 2;

QY 3 CINQLLCKLPTDSELA PRSXCC-HCRH 28
DB 38 CYNKCYCK-----RCCYHCOH 53

RESULT 40
TAT_SIVSP
ID TAT_SIVSP STANDARD; PRT; 128 AA.
AC P19507;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TAT protein (Transactivating regulatory protein).
GN TAT
OS Simian immunodeficiency virus (PBj14/BCL-3 isolate) (sooty mangabey).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11738;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90272009; PubMed=1971917;
RA Dewhurst S., Embretson J.E., Anderson D.C., Mullins J.I., Fultz P.N.;
RT "Sequence analysis and acute pathogenicity of molecularly cloned
RT SIVSMM-PBj14.";
RL Nature 345:636-640(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92368737; PubMed=1503826;
RA Dewhurst S., Embretson J.E., Fultz P.N., Mullins J.I.;
RT "Molecular clones from a non-acutely pathogenic derivative of
RT SIVsmPBj14: characterization and comparison to acutely pathogenic
RT clones.";
RL AIDS Res. Hum. Retroviruses 8:1179-1187(1992).
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC
CC -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; L03295; AAB59768.1; -
DR EMBL; L03298; AAA47775.1; -
DR EMBL; M31325; AAA47750.1; -
DR HIV; M31325; TATSSMMPBj.
DR InterPro; IPR001831; HIV_Tat.
DR Pfam; PF00539; Tat; 1.
DR PRINTS; PR00055; HIVTATDOMAIN.
KW Transcription regulation; Activator; RNA-binding; Nuclear protein;
KW AIDS.
SQ SEQUENCE 128 AA; 14582 MW; 9DC6A467927D4403 CRC64;

Query Match 26.8%; Score 44; DB 1; Length 128;
Best Local Similarity 33.3%; Pred. No. 19;
Matches 9; Conservative 2; Mismatches 4; Indels 12; Gaps 2;

QY 3 CINQLLCKLPTDSELA PRSXCC-HCRH 28
DB 50 CYNKCYCK-----RCCYHCOH 65

Search completed: August 22, 2002, 07:52:31

Job time: 189 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 07:48:17 ; Search time 27.12 Seconds
(without alignments)
102.750 Million cell updates/sec

Title: US-09-856-199-1
Perfect score: 184
Sequence: 1 KRCINQLLCKLPTDSELA PRSXCHCRHL 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	161	98.2	381	1 A47327	selenoprotein p pr
2	122	74.4	380	2 T10442	selenoprotein p pr
3	107	65.2	385	1 OMRTSP	selenoprotein p pr
4	53	32.3	324	2 C86582	hypothetical prote
5	53	32.3	324	2 D72041	conserved hypothet
6	53	32.3	1662	1 H71402	probable kinesin -
7	51.5	31.4	518	2 T33196	hypothetical prote
8	51.5	31.4	521	1 VGIVDH	envelope glycoprot
9	51	31.1	534	2 T22154	hypothetical prote
10	51	31.1	556	2 T45066	steroid hormone re
11	50.5	30.8	772	2 H84605	hypothetical prote
12	50	30.5	229	2 AG2171	hypothetical prote
13	50	30.5	588	2 S34685	probable purine nu
14	50	30.5	617	2 S48160	metalloproteinase
15	49.5	30.2	248	1 A33490	type IV prepillin p
16	49.5	30.2	1009	2 S64734	retrovirus-related
17	49	28.9	754	2 D70536	probable purL prot
18	48	28.3	290	1 S70875	type IV prepillin p
19	48	29.3	313	2 T49129	hypothetical prote
20	48	29.3	558	2 T10672	aspartate--trNA li
21	48	29.3	1042	2 T26644	hypothetical prote
22	48	29.3	1268	2 B36502	insulin receptor-r
23	47	28.7	83	2 F35982	bitan alpha - puff
24	47	28.7	197	2 E97930	conserved hypothet
25	47	28.7	228	2 T20219	hypothetical prote
26	47	28.7	591	2 G86445	hypothetical prote
27	47	28.7	659	2 A64139	ribonuclease T2 ho
28	47	28.7	789	2 A48337	genome polyprotein
29	47	28.7	1286	1 S38058	hypothetical prote

30	47	28.7	1722	2 E89753	protein FilC7.4 li
31	46.5	28.4	383	2 F69018	formate dehydrogen
32	46	28.0	119	2 A46356	tat protein - siml
33	46	28.0	324	2 F22848	hypothetical ORF-4
34	46	28.0	344	2 C25877	hypothetical prote
35	46	28.0	507	1 KJH0AA	cerbroside-sulfat
36	46	28.0	553	2 B69167	glutamate--trNA li
37	46	28.0	736	2 T40510	beta transducin -
38	45.5	27.7	1801	1 MMRTS	laminin beta-2 cha
39	45	27.4	172	2 S24397	stellate protein -
40	45	27.4	175	2 C71415	hypothetical prote
41	45	27.4	188	2 AF3559	hypothetical prote
42	45	27.4	434	2 F84605	hypothetical prote
43	45	27.4	632	2 S58152	hypothetical prote
44	45	27.4	826	2 A60385	monocyte surface a
45	45	27.4	902	2 T47966	hypothetical prote
46	45	27.4	1193	2 A86193	hypothetical prote
47	45	27.4	2150	2 T32497	hypothetical prote
48	45	27.4	3229	2 S27852	probable cell-surf
49	44.5	27.1	312	2 T32379	hypothetical prote
50	44.5	27.1	591	2 I48141	acroganin - guine
51	44.5	27.1	1391	2 T20406	hypothetical prote
52	44	26.8	63	2 S08191	metallothionein 2
53	44	26.8	112	2 A46717	colipase precursor
54	44	26.8	112	2 S62929	hypothetical prote
55	44	26.8	131	2 T11557	tat protein - siml
56	44	26.8	132	2 T11564	tat protein - siml
57	44	26.8	182	2 T34009	hypothetical prote
58	44	26.8	247	2 T30393	hypothetical prote
59	44	26.8	249	2 F83477	hypothetical prote
60	44	26.8	311	2 T15268	hypothetical prote
61	44	26.8	388	1 WMV202	major envelope ant
62	44	26.8	1338	2 A49634	aldehyde oxidase (
63	44	26.8	2167	2 S50658	bud emergence prot
64	43.5	26.5	86	2 D91119	hypothetical prote
65	43.5	26.5	86	2 S11458	hypothetical 9.9K
66	43.5	26.5	86	2 C85964	hypothetical prote
67	43.5	26.5	183	1 QQBEC7	HXLf6 protein prec
68	43.5	26.5	269	1 S74490	type IV prepillin p
69	43.5	26.5	308	2 B83403	ribokinase PA1950
70	43.5	26.5	775	2 E84828	probable WD-40 rep
71	43.5	26.5	884	2 T18649	hypothetical prote
72	43	26.2	63	2 C34620	metallothionein -
73	43	26.2	63	2 S33381	metallothionein -
74	43	26.2	63	2 A34620	metallothionein -
75	43	26.2	63	2 A34958	metallothionein -
76	43	26.2	75	2 JX0169	cytotoxic factor I
77	43	26.2	97	2 T18173	hypothetical prote
78	43	26.2	113	2 B72629	hypothetical prote
79	43	26.2	175	2 B83411	hypothetical prote
80	43	26.2	307	2 S51373	tau-protein kinase
81	43	26.2	339	2 S37920	MBR1 protein precu
82	43	26.2	353	2 AG0138	quinolate synthet
83	43	26.2	378	2 T00481	probable RING zinc
84	43	26.2	494	2 T09289	nicotinic acetylch
85	43	26.2	571	1 RNCW7T	transcription init
86	43	26.2	571	1 RNCW7H	transcription init
87	43	26.2	571	2 I40872	transcription init
88	43	26.2	571	2 A81653	RNA polymerase sig
89	43	26.2	572	2 E72039	RNA polymerase sig
90	43	26.2	572	2 B86585	RNA polymerase sig
91	43	26.2	591	2 E81502	RNA polymerase sig
92	43	26.2	754	2 F87185	phosphoribosylform
93	43	26.2	908	2 H86537	oxoglutarate dehyd
94	43	26.2	908	2 G72085	2-oxoglutarate deh
95	43	26.2	1261	2 S75130	sensory transducti
96	43	26.2	1297	2 T30274	proteoliasin - se
97	43	26.2	1404	1 A48196	protein-tyrosine k
98	43	26.2	1440	2 T27942	lin-15B protein -
99	42.5	25.9	167	2 S02515	nifO protein - Kie
100	42.5	25.9	236	2 D81652	coenzyme PQQ synth

ALIGNMENTS

RESULT 1
A47327
selenoprotein P precursor [validated] - human
C:Species: Homo sapiens (man)
C:Date: 21-Sep-1993 #sequence_revision 01-Dec-1995 #text_change 15-Sep-2000
C:Accession: A47327; S42752
R: Hill, K.E.; Lloyd, R.S.; Burk, R.F.
Proc. Natl. Acad. Sci. U.S.A. 90, 537-541, 1993
A:Title: Conserved nucleotide sequences in the open reading frame and 3' untranslated region
A:Reference number: A47327; MUID:93133823
A:Accession: A47327
A:Molecule type: mRNA
A:Residues: 1-381 <HIL>
A:Cross-references: GB:211793; NID:g36425; PIDN:CAA77836.1; PID:g2654365
A:Experimental source: heart and liver
A:Note: In Genbank entry HSSSLPM, release 117.0, PIDN:CAA77836.1, the selenocysteine UGA
R:Akesson, B.; Billew, T.; Burk, R.F.
Biochim. Biophys. Acta 1204, 243-249, 1994
A:Title: Purification of selenoprotein P from human plasma.
A:Reference number: S42752; MUID:94191007
A:Accession: S42752
A:Molecule type: protein
A:Residues: 20-27, 'X', 29-33 <AKE>
A:Note: mature forms of 55K and 61K were detected in plasma; the protein was shown to co
C:Genetics:
A:Gene: GDB:SEP1; SLNP
A:Cross-references: GDB:138278; OMIM:601484
A:Map position: 5q31-5q31
C:Function:
A:Description: may act as a free-radical scavenger
C:Superfamily: selenoprotein P
C:Keywords: extracellular protein; glycoprotein; heparin binding; liver; plasma; seleno
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-381/Product: selenoprotein P #status experimental <MAT>
F:46,83,119,128,338/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:59,300,318,330,345,352,367,369,376,378/Modified site: selenocysteine #status predicted

Query Match 98.2%; Score 161; DB 1; Length 381;
Best Local Similarity 96.6%; Pred. No. 2.2e-14;
Matches 28; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KRCINOLLCKLPDSEAPRSXCCHCRHL 29
|||||
DB 279 KRCINOLLCKLPDSEAPRSXCCHCRHL 307

RESULT 2
T10442
selenoprotein P precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T10442
R:Steinert, P.; Ahrens, M.; Gross, G.; Flohe, L.
Biofactors 6, 311-319, 1997
A:Title: cDNA and deduced polypeptide sequence of a mouse selenoprotein P.
A:Reference number: Z17017; MUID:97434516
A:Accession: T10442
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-380 <STE>
A:Cross-references: EMBL:X99807; NID:g1495998; PIDN:CAA68140.1; PID:g1495999
C:Genetics:
A:Gene: selP
C:Superfamily: selenoprotein P
C:Keywords: extracellular protein; selenocysteine

Query Match 74.4%; Score 122; DB 2; Length 380;
Best Local Similarity 77.8%; Pred. No. 4.4e-09;

Matches 21; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 CINOLLCKLPDSEAPRSXCCHCRHL 29
|||||
DB 281 CINOLLCKLSEAPRSXCCHCRHL 307

RESULT 3
OMRTSP
selenoprotein P precursor [validated] - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 15-Sep-2000
C:Accession: A40380; B40380; S68322
R:Hill, K.E.; Lloyd, R.S.; Yang, J.G.; Read, R.; Burk, R.F.
J. Biol. Chem. 266, 10050-10053, 1991
A:Title: The cDNA for rat selenoprotein P contains 10 TGA codons in the open reading
A:Reference number: A40380; MUID:91244760
A:Accession: A40380
A:Molecule type: mRNA
A:Residues: 1-385 <HIL>
A:Cross-references: GB:M63574; NID:g206893; PIDN:AAA42129.1; PID:g206894
A:Accession: B40380
A:Molecule type: protein
A:Residues: 20-41; 267-287; 316-327 <HI2>
R:Chittum, H.S.; Himeno, S.; Hill, K.E.; Burk, R.F.
Arch. Biochem. Biophys. 325, 124-128, 1996
A:Title: Multiple forms of selenoprotein P in rat plasma.
A:Reference number: S68322; MUID:96140605
A:Accession: S68322
A:Molecule type: protein
A:Residues: 20-27 <CHI>
C:Superfamily: selenoprotein P
C:Keywords: extracellular protein; glycoprotein; heparin binding; liver; plasma; sele
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-385/Product: selenoprotein P #status experimental <MAT>
F:59,264,335,357,371,373,380,382/Modified site: selenocysteine #status predicted
F:83,174,188,370,375/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:282,323/Modified site: selenocysteine #status experimental

Query Match 65.2%; Score 107; DB 1; Length 385;
Best Local Similarity 70.4%; Pred. No. 4.8e-07;
Matches 19; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 3 CINOLLCKLPDSEAPRSXCCHCRHL 29
|||||
DB 286 CINOLLCKLSESGAATSCCCHCRHL 312

RESULT 4
C86582
hypothetical protein ycea [imported] - Chlamydomonas pneumoniae (strain J138)
C:Species: Chlamydomonas pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: C86582
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349
A:Accession: C86582
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-324 <STO>
A:Cross-references: GB:BA000008; NID:g9879106; PIDN:BA098941.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: ycea
C:Superfamily: Escherichia coli ycea protein

Query Match 32.3%; Score 53; DB 2; Length 324;
Best Local Similarity 46.7%; Pred. No. 8.9;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 13 TDSEAPRSXCCHCR 27
: : : : :
Db 236 SDPDVAPIAECCHCQ 250

RESULT 5
D72041
conserved hypothetical protein CP0012 [imported] - Chlamydomophila pneumoniae (strains CWI
N:Alternate names: ycea hypothetical protein
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: D72041; F81622
R:Kalan, S.; Mitchell, W.; Marathe, R.; Lammell, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative Genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606
A:Accession: D72041
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-324 <ARN>
A:Cross-references: GB:AE001363; NID:g4377039; PIDN:AAI8873.1; PID:g4377704
A:Experimental source: strain CWL029
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255
A:Accession: F81622
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-324 <REA>
A:Cross-references: GB:AE002165; GB:AE002161; NID:g7188948; PIDN:AAF37908.1; PID:g718895
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: ycea; CP0012
C:Superfamily: Escherichia coli ycea protein

Query Match 32.3%; Score 53; DB 2; Length 324;
Best Local Similarity 46.7%; Pred. No. 8.9;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 13 TDSEAPRSXCCHCR 27
: : : : :
Db 236 SDPDVAPIAECCHCQ 250

RESULT 6
H71402
probable kinesin - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
A:Variety: Columbia
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: H71402
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terry, N.; Giel
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans
erhof, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans
C.; Chalwatzis, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A:Reference number: A71400; MUID:98121113
A:Accession: H71402
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1662 <BEV>
A:Cross-references: GB:Z97335; NID:g2244747; PIDN:CAB10194.1; PID:g2244771
C:Genetics:
A:Map position: 4COP9-4G3845
C:Superfamily: Arabidopsis thaliana 186.9K kinesin-related protein; kinesin motor domain
F;43-423/Domain: kinesin motor domain homology <RMOT>

Query Match 32.3%; Score 53; DB 1; Length 1662;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 KLPTDSEAPRSXCCH 25
: : : : :
Db 543 RLPEDSDVAMEDACCH 558

RESULT 7
T33196
hypothetical protein F36F12.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 26-May-2000
C:Accession: T33196
R:Smith, A.; Harmon, G.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of C. elegans cosmid F36F12.
A:Reference number: Z21300
A:Accession: T33196
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-518 <SMI>
A:Cross-references: EMBL:AF067940; PIDN:AAC19200.1; GSPDB:GN00023; CBSP:F36F12.1
A:Experimental source: strain Bristol N2; clone F36F12
C:Genetics:
A:Gene: CESP:F36F12.1
A:Map position: 5
A:Introns: 74/2; 118/2; 154/3; 315/1; 394/3; 430/2; 475/3
C:Superfamily: Caenorhabditis elegans hypothetical protein F07G11.3

Query Match 31.4%; Score 51.5; DB 2; Length 518;
Best Local Similarity 33.3%; Pred. No. 21;
Matches 9; Conservative 5; Mismatches 6; Indels 7; Gaps 1;

QY 2 RCINOLLCKLPDSEAPRSXCCHCRH 28
: : : : :
Db 467 QCVNTRKACKLPNDKIP-----CRH 486

RESULT 8
VGIVDH
envelope glycoprotein - Dhori virus (strain India/1313/61)
C:Species: Dhori virus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jul-1999
C:Accession: A34679
R:Freedman-Faulstich, E.Z.; Fuller, F.J.
Virology 175, 10-18, 1990
A:Title: Nucleotide sequence of the tick-borne, orthomyxo-like Dhori/Indian/1313/61 v
A:Reference number: A34679; MUID:90177204
A:Accession: A34679
A:Molecule type: mRNA
A:Residues: 1-521 <PRE>
A:Cross-references: EMBL:M34002; NID:g335189; PIDN:AAA47907.1; PID:g335190
C:Genetics:
A:Map position: segment 4
C:Superfamily: baculovirus major envelope glycoprotein
C:Keywords: envelope protein; glycoprotein; transmembrane protein
F;44,158,189,396/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 31.4%; Score 51.5; DB 1; Length 521;
Best Local Similarity 44.0%; Pred. No. 21;
Matches 11; Conservative 3; Mismatches 8; Indels 3; Gaps 1;

QY 3 CINOLLCKLPDSELP-----APRSXCCH 24
: : : : :
Db 88 CYNDLIPKSPTESELRWKSQKCC 112

RESULT 9

T22154
hypothetical protein F44A6.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T22154
R: Sulston, J.
submitted to the EMBL Data Library, August 1995
A:Reference number: Z19524
A:Accession: T22154
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-534 <WIL>
A:Cross-references: EMBL:Z50858; PIDN:CAA90722.1; GSPDB:GN00028; CESP:F44A6.2
A:Experimental source: clone F44A6
C:Genetics:
A:Gene: CESP:F44A6.2
A:Map position: X
A:Introns: 58/3; 123/1; 172/3; 220/1; 285/3; 464/1; 506/3
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology

Query Match 31.1%; Score 51; DB 2; Length 534;
Best Local Similarity 34.6%; Pred. No. 25;
Matches 9; Conservative 7; Mismatches 6; Indels 4; Gaps 1;

QY 2 RCINQLCKLPTDSELAPRSXCCHCR 27

||:|:|::|:|:|:|

Db 186 RCLKQVCEIKRES----RNRCOYCR 207

RESULT 10

I45066
steroid hormone receptor CNR14 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 09-Mar-1996 #sequence_revision 09-Mar-1996 #text_change 21-Jul-2000
C:Accession: I45066; T42757
R: Kostrouch, Z.; Kostrouchova, M.; Rall, J.E.
Proc. Natl. Acad. Sci. U.S.A. 92, 156-159, 1995
A:Title: Steroid/thyroid hormone receptor genes in Caenorhabditis elegans.
A:Reference number: A55595; MUID:95116514
A:Accession: T42757
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-556 <RES>
A:Cross-references: EMBL:U13074; NID:g538368; PID:g538369
R: Carmi, I.; Kopczyński, J.B.; Meyer, B.J.
Nature 396, 168-173, 1998
A:Title: The nuclear hormone receptor SEX-1 is an X-chromosome signal that determines ne
A:Reference number: Z22264; MUID:99039502
A:Accession: T42757
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-536 <CAR>
A:Cross-references: EMBL:AF049869; NID:g3882947; PIDN:AAC77795.1; PID:g3882948
C:Genetics:
A:Gene: cnr14; sex-1
A:Map position: X
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
C:Keywords: steroid hormone receptor; zinc finger
F:149-408/Domain: erba transforming protein homology <ERBA>

Query Match 31.1%; Score 51; DB 2; Length 556;
Best Local Similarity 34.6%; Pred. No. 26;
Matches 9; Conservative 7; Mismatches 6; Indels 4; Gaps 1;

QY 2 RCINQLCKLPTDSELAPRSXCCHCR 27

||:|:|::|:|:|:|

Db 186 RCLKQVCEIKRES----RNRCOYCR 207

RESULT 11

H84605

hypothetical protein At2g21850 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: H84605
R: Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: H84605
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-772 <STO>
A:Cross-references: GB:AE002093; NID:g4417297; PIDN:AAD20422.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g21850
A:Map position: 2

Query Match 30.8%; Score 50.5; DB 2; Length 772;
Best Local Similarity 40.0%; Pred. No. 40;
Matches 14; Conservative 3; Mismatches 7; Indels 11; Gaps 2;

QY 1 KRCI-----NQLLCKLPTDSELAPRSXCCHCRH 28

||||:|:|:|:|

Db 131 KRCPGPDARGSCLLCELP----LSPSSVCYGCVV 161

RESULT 12

AG2171
hypothetical protein alr2926 [imported] - Anabaena sp. (strain PCC 7120)
C:Species: Anabaena sp.
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C:Accession: AG2171
R: Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG2171
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-229 <KUR>
A:Cross-references: GB:BA000019; PIDN:BAB74625.1; PID:gl7132020; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr2926

Query Match 30.5%; Score 50; DB 2; Length 229;
Best Local Similarity 40.7%; Pred. No. 17;
Matches 11; Conservative 3; Mismatches 5; Indels 8; Gaps 2;

QY 8 LCKLPTDSELAPRSXC-----CHCRH 28

||:|:|:|:|:|:|

Db 22 LCQRTSTELCPN--CTROLQKCHHKH 46

RESULT 13

S34685
probable purine nucleotide-binding protein YKL185w - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 06-Feb-1998
C:Accession: S34685; S38017; S27436
R: Wieman, S.; Voss, H.; Schwagaer, C.; Rupp, T.; Stegemann, J.; Zimmermann, J.; Groth
submitted to the EMBL Data Library, July 1993
A:Description: Sequencing and analysis of 51.5 kilobases on the left arm of chromosome
A:Reference number: S34679
A:Accession: S34685
A:Molecule type: DNA

A:Title: Molecular structure of a gypsy element of Drosophila subobscura (gypsyds) const
A:Reference number: S64733; MUID:96174659
A:Accession: S64734
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-1009 <ALB>
A:Cross-references: EMBL:X72390
A:Experimental source: strain H271
R:Alberola, T.M.
submitted to the EMBL Data Library, March 1993
A:Reference number: S72396
A:Accession: S72396
A:Molecule type: DNA
A:Residues: 1-236, '1', 238-1009 <ALW>
A:Cross-references: EMBL:X72390; NID:g1237203; PIDN:CAA51084.1; PID:g1237205
R:Alberola, T.M.; de Frutos, R.
J. Mol. Evol. 36, 127-135, 1993
A:Title: Gypsy homologous sequences in Drosophila subobscura (gypsyds).
A:Reference number: S33116; MUID:93164273
A:Accession: S33116
A:Molecule type: DNA
A:Residues: 381-815 <ALF>
A:Cross-references: EMBL:X63724
A:Experimental source: strain H271
C:Genetics:
A:Mobile element: retrotransposon gypsy
C:Superfamily: pol polyprotein
C:Keywords: polyprotein; reverse transcriptase
F:4/Active site: Asp (shared with dimeric partner) #status predicted

Query Match 30.2%; Score 49.5; DB 2; Length 1009;
Best Local Similarity 33.3%; Pred. No. 69;
Matches 11; Conservative 6; Mismatches 9; Indels 7; Gaps 1;

QY 4 INQLLCKLPTDSE-----LAPRSXCCHRH 29
DB 670 VNAILCDLPTLASQHDLLILFFATQFWHCKNL 702

RESULT 17
D70536
Probable purL protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: D70536
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987
A:Accession: D70536
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-754 <COL>
A:Cross-references: GB:D295618; GB:AL123456; NID:g3261788; PIDN:CAB09108.1; PID:g3261789
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: purL
C:Superfamily: phosphoribosylformylglycinamide synthase component II

Query Match 29.9%; Score 49; DB 2; Length 754;
Best Local Similarity 44.8%; Pred. No. 63;
Matches 13; Conservative 4; Mismatches 10; Indels 2; Gaps 1;

QY 2 RCINQLLCKLPTDSELA PRSXC--HCRH 28
DB 29 RRIRQILGRPTDTLAMYVMWSEHCYS 57

RESULT 18
S70875
type IV prepilin peptidase (EC 3.4.99.-) tapD - Aeromonas hydrophila
N:Alternate names: prepilin type IV peptidase
N:Contains: type IV pilin N-methyltransferase (EC 2.1.1.-)
C:Species: Aeromonas hydrophila
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S70875
R:Pepe, C.M.; Eklund, M.W.; Strom, M.S.
Mol. Microbiol. 19, 857-869, 1996
A:Title: Cloning of an Aeromonas hydrophila type IV pilus biogenesis gene cluster: co-
quired for extracellular protein secretion.
A:Reference number: S70872; MUID:96417863
A:Accession: S70875
A:Molecule type: DNA
A:Residues: 1-230 <PEP>
A:Cross-references: EMBL:U20255; NID:g695159; PIDN:AAC43998.1; PID:g6631118
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995
C:Genetics:
A:Gene: tapD
C:Superfamily: type IV prepilin peptidase
C:Keywords: hydrolase; methyltransferase; S-adenosylmethionine

Query Match 29.3%; Score 48; DB 1; Length 290;
Best Local Similarity 47.1%; Pred. No. 39;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 12 PTDSELA PRSXCCHCRH 28
DB 63 PYNLMVPRSCCPHCNH 79

RESULT 19
T49129
hypothetical protein F26G5.80 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 24-Aug-2001
C:Accession: T49129
R:D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Rudd, S.
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z25017
A:Accession: T49129
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-313 <DAN>
A:Cross-references: EMBL:AL353814; GSPDB:GN00061; ATSP:F26G5.80
A:Experimental source: cultivar Columbia; BAC clone F26G5
C:Genetics:
A:Gene: ATSP:F26G5.80
A:Map position: 3
A:Introns: 266/1; 303/3
C:Superfamily: Arabidopsis thaliana hypothetical protein F24M12.210

Query Match 29.3%; Score 48; DB 2; Length 313;
Best Local Similarity 45.0%; Pred. No. 41;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 6 QLCLKLPTDSELA PRSXCCH 25
DB 13 EILCLPLGLSLVRLSVCKH 32

RESULT 20
T10672
aspartate--tRNA ligase homolog F6E21.100 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 24-Nov-1999
C:Accession: T10672
R:Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Ban
submitted to the Protein Sequence Database, June 1999

C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T20219; T21470
R:Dobson, R.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19238
A:Accession: T20219
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-228 <W12>
A:Cross-references: EMBL:Z75531; PIDN:CAA99808.1; GSPDB:GN00023; CESP:C54D10.10
A:Experimental source: clone C54D10
R:Kershaw, J.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19427
A:Accession: T21470
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-228 <W12>
A:Cross-references: EMBL:Z75539; PIDN:CAA99845.1; GSPDB:GN00023; CESP:C54D10.10
A:Experimental source: clone F28C1
C:Genetics:
A:Gene: CESP:C54D10.10
A:Map position: 5
A:Introns: 11/1; 28/1; 59/1; 130/1

Query Match 28.7%; Score 47; DB 2; Length 228;
Best Local Similarity 43.8%; Pred. No. 44;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 9 CKLPTDSELAPRSXCC 24
|||||::: :|
Db 133 CKLPTDAKIVTKAOKC 148

RESULT 26
G86445
Hypothetical protein F3C3.15 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C:Accession: G86445
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: G86445
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-591 <STO>
A:Cross-references: GB:AE005172; NID:g10801368; PIDN:AAG23440.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 28.7%; Score 47; DB 2; Length 591;
Best Local Similarity 40.0%; Pred. No. 96;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 4 INQLCKLPTDSELAPRSXC 23
: :|||::| | ||| |
Db 12 VEEILCRVPTSLKAVRSC 31

RESULT 27
A64139

ribonuclease T2 homolog - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 21-Jul-2000
C:Accession: A64139
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630
A:Accession: A64139
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-659 <TIGR>
A:Cross-references: GB:U32846; GB:L42023; NID:g3212237; PIDN:AAC23378.1; PID:g1574590
A:Note: named as homolog to a protein from Escherichia coli
C:Superfamily: exoribonuclease II

Query Match 28.7%; Score 47; DB 2; Length 659;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 10; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

QY 4 INQLCKLPTDSELAP 19
|||||:| | ||||
Db 147 INQLICR--ADDELAP 160

RESULT 28
A48337
genome polyprotein - potato virus A (fragment)
C:Species: potato virus A, PVA
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 17-Nov-2000
C:Accession: A48337
R:Collins, R.F.; Leclerc, D.; AbouHaidar, M.G.
Arch. Virol. 128, 135-142, 1993
A:Title: Cloning and nucleotide sequence of the capsid protein and the nuclear inclus
A:Reference number: A48337; MUID:93119266
A:Accession: A48337
A:Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-789 <COL>
A:Cross-references: GB:S51667; NID:g262857; PIDN:AAB24779.1; PID:g262858
C:Note: sequence extracted from NCBI backbone (NCBI:121715, NCBI:121716)
C:Superfamily: tobacco etch virus genome polyprotein
C:Keywords: polyprotein

Query Match 28.7%; Score 47; DB 2; Length 789;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 NQLCKLPTD 14
|||:||||| |
Db 236 NQLMCKLPDD 245

RESULT 29
S38058
Hypothetical protein YKL215c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein F1286
C:Species: Saccharomyces cerevisiae
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C:Accession: S38058; S38053; S44323; S17009
R:Alexandraki, D.; Tzermia, M.
submitted to the Protein Sequence Database, March 1994
A:Reference number: S38054
A:Accession: S38058
A:Molecule type: DNA
A:Residues: 1-1286 <ALE>
A:Cross-references: EMBL:Z28215; NID:g486384; PID:g486385; GSPDB:GN00011; MIPS:YKL215
A:Experimental source: strain S288C

A;Accession: F69018
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-383 <MTH>
A;Cross-references: GB:AE000884; GB:AE000666; NID:G2622242; PIDN:AAB85628.1; PID:G2662;
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH139
A;Start codon: GTG
C;Superfamily: formate dehydrogenase chain B: ferredoxin [2Fe-4S] homology

```

Query Match      28.48; Score 46.5; DB 2; Length 383;
Best Local Similarity 38.5%; Pred. No. 78;
Matches 10; Conservative 2; Mismatches 9; Indels 5; Gaps 1;

Qy 1 KRCINQLLCKLPDSELA PRSXKCHC 26
      ||||| : ||| |
Db 281 KRCINCFACR-----DLCPICFCRC 301

```

RESULT 32
A46356
tat protein - simian immunodeficiency virus SIVagm (type 3)
C:Species: simian immunodeficiency virus SIVagm
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 20-Sep-1999
C:Accession: A46356; A36860
R:Baier, M.; Garber, C.; Mueller, C.; Cichutek, K.; Kurth, R.

A: r1c1e: Complete nucleotide sequence of a simian immunodeficiency virus from Africa
A: Reference number: A46356; MUID: 90232731
A: Accession: A46356
A: Status: nucleic acid sequence not shown
A: Molecule type: DNA
A: Residues: 1-119 <BAI>
A: Cross-references: GB:M30931; NID:G334400; PIDN:AAA91917.1; PID:G334401
C: Genetics:
C: Introns: 73/2
C: Superfamily: AIDS trans-activating transcription regulator
C: Keywords: transcription regulation

hypothetical ORF-4 protein - Leishmania tarentolae mitochondrion
C:Species: mitochondrion Leishmania tarentolae
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 07-Dec-1999
C:Accession: F22848
R:de la Cruz, V.F.; Neckelmann, N.; Simpson, L.
J. Biol. Chem. 259, 15136-15147, 1984
A:Title: Sequences of six genes and several open reading frames in the kinetoplast ma
A:Reference number: A22848; MUID:85079995

A;Accession: F22848
A;Reference number: A22848; MUID:850/9995
A:Molecule type: DNA

A:Residues: 1-324
A:Cross-references: GB:M10126
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC6
C:Superfamily: NADH dehydrogenase
C:Keywords: mitochondrion

Query Match 28.0%; Score 46; DB 2; Length 324;
Best Local Similarity 31.4%; Pred. No. 80;
Matches 16; Conservative 1; Mismatches 6; Indels

RESULT 34

C25877

hypothetical protein 1 - Crithidia fasciculata mitochondrion

C:Species: Mitochondrion Crithidia fasciculata

C:Date: 19-Jan-1988 #sequence_revision 19-Jan-1988 #text_change 07-Dec-1999

C:Accession: C25877

R:Sloof, P.; van den Burg, J.; Voogd, A.; Benne, R.

Nucleic Acids Res. 15, 51-65, 1987

A:Title: The nucleotide sequence of a 3.2 kb segment of mitochondrial maxicircle DNA from Crithidia fasciculata. The sequence encodes a 1080 base pair open reading frame for a 360 amino acid protein, which is homologous to the cytochrome b gene and a possible frameshift gene; further evidence for the use of unusual start codons in the mitochondrial genome of Crithidia fasciculata.

A:Reference number: A25877; MUID:87146364

A:Accession: C25877

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-344 <SLO>

C:Genetics:

A:Genome: mitochondrion

A:Genetic code: SGC6

C:Superfamily: NADH dehydrogenase (ubiquinone) 49K protein

C:Keywords: mitochondrion

Query Match 28.0%; Score 46; DB 2; Length 344;
Best Local Similarity 31.4%; Pred. No. 84;
Matches 16; Conservative 1; Mismatches 6; Indels 2

RESULT 35
KJHUAA
cerebroside-sulfatase [EC 3.1.6.8] precursor - human
N; Alternate names: arylsulfatase (EC 3.1.6.1) [misidentification]; arylsulfatase A (ASA)
C; Species: Homo sapiens (man)
C; Date: 21-Nov-1993 #sequence_revision 27-Oct-1995 #text_change 04-Feb-2000
C; Accession: S11031; G02857; A32207; S23932
R; Kreysing, J.; von Figura, K.; Gieselmann, V.
Eur. J. Biochem. 191, 627-631, 1990
A; Title: Structure of the arylsulfatase A gene.
A; Reference number: S11031; MUID:90361046
A; Accession: S11031
A; Molecule type: DNA
A; Residues: 1-507 <KRE>
A; Cross-references: EMBL:X52150; NID:g28859; PIDN:CAA36398.1; PID:g28860
R; Adams, M.D.; Kerlavage, A.R.; Fuldner, R.A.; Phillips, C.A.; Venter, J.C.
submitted to the EMBL Data Library, June 1996
A; Reference number: H01749
A; Accession: G02857
A; Status: translated from GB/EMBL/DDBJ
A; Molecule type: DNA
A; Residues: 'MS', 1-507 <ADA>
A; Cross-references: EMBL:U62317; NID:g1399959; PID:g13
A; Note: an incorrect initiation codon was used

R;Stein, C.; Gieselmann, V.; Kreysing, J.; Schmidt, B.; Pohlmann, R.; Waheed, A.; Mey J. Biol. Chem. 264, 1252-1259, 1989

A;Title: Cloning and expression of human arylsulfatase A.

A;Reference number: A32207; MUID:89093115

A;Accession: A32207

A;Molecule type: mRNA

A;Residues: 1-358; 'RPPAAGHRQEPSAVSLLPLVLRPGPWGFCADMKVQGSLLHP', 402-507 <STE>

A;Cross-references: GB:X52151; GB:J04442; GB:J04593

A;Title: parts of this sequence, including the amino end of the mature protein, were d

R;Nujii, T.; Kobayashi, T.; Honke, K.; Gasa, S.; Ishikawa, M.; Shimizu, T.; Makita, A. Biochim. Biophys. Acta 1122, 93-98, 1992

A;Title: Proteolytic processing of human lysosomal arylsulfatase A.

A;Reference number: S29932; MUID:92338230

A;Accession: S29932

A;Molecule type: protein

A;Residues: 20-29, 31-33; 434-479 <PUJ>

A;Experimental source: Placenta

R;Schmidt, B.; Selmer, T.; Ingendoh, A.; von Figura, K. Cell 82, 271-278, 1995

A;Title: A novel amino acid modification in sulfatases that is defective in multiple

A;Reference number: A57113; MUID:95354208

A;Contents: annotation; identification of 3-oxoalanine, 2-amino-3-oxopropanoic acid

C;Genetics:

A;Gene: GDB:ARSA

A;Cross-references: GDB:119007; OMIM:250100

A;Map position: 22q13.31-22qter

A;Introns: 73/2; 153/3; 226/3; 283/2; 325/1; 367/3; 402/1

A;Note: defects in this gene can cause metachromatic leukodystrophy

C;Function:

A;Description: hydrolyzes cerebroside 3-sulfate to release sulfate; can also hydrolyz

C;Superfamily: animal sulfatase

C;Keywords: glycoprotein; lysosomal storage disease; lysosome; sulfuric ester hydrola

F;1-19/Domain: signal sequence #status predicted <SIG>

F;20-507/Product: cerebroside-sulfatase #status predicted <MAT>

F;20-444, 445-507/Product: cerebroside-sulfatase component b #status predicted <MCB>

F;20-444, 448-507/Product: cerebroside-sulfatase component c #status predicted <MCC>

F;20-444, 449-507/Product: cerebroside-sulfatase minor component c #status predicted <MCC>

F;69/Modified site: 3-oxoalanine (Cys) #status experimental

F;158/Binding site: carbohydrate (Asn) (covalent) #status experimental

F;158,350/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	28.0%;	Score 46;	DB 1;	Length 507;
Best Local Similarity	75.0%;	Pred. No. 1.2e+02;		
Matches	6;	Conservative	0;	Mismatches
			2;	Indels
			0;	Gaps
Oy	19	PRSXCHC 26		
Db	495	PRPACCHC 502		
RESULT	36			
B69167				
		glutamate--tRNA ligase (EC 6.1.1.17) - Methanobacterium thermoautotrophicum (strain D		
N;Alternate names:		glutamyl-tRNA synthetase		
C;Species:		Methanobacterium thermoautotrophicum		
C;Date:		05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 16-Jul-1999		
C;Accession:		B69167		
R;Smith, D.R.;		Doucette-Stamm, L.A.;		
		Deloughery, C.;		
		Lee, H.;		
		Dubois, J.;		
		Aldredge, T.;		
		Qlu, D.;		
		Spadafora, R.;		
		Vicaire, R.;		
		Wang, Y.;		
		Wierzbowski, J.;		
		Gibson, R.;		
		Jiwani, k.;		
		S.;		
		Church, G.M.;		
		Daniels, C.J.;		
		Mao, J.;		
		Rice, P.;		
		Noelling, J.;		
		Reeve, J.N.		
J. Bacteriol.	179,	7135-7155,	1997	
A;Title:		Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu		
A;Reference number:		A69000;MUID:98037514		

C:Superfamily: glutamine--tRNA ligase; glutamine--tRNA ligase homology
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis
F:97-371/Domain: glutamine-tRNA ligase homology <EGL>

Query Match 28.0%; Score 46; DB 2; Length 553;
Best Local Similarity 42.9%; Pred. No. 1.2e+02;
Matches 9; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 9 CKLPTDSELA PRSXCHCRHL 29
II II IIIII

Db 193 CRPEEFRELKNRGEACHCRSL 213

RESULT 37

T40510 beta transducin - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 26-May-2000
C:Accession: T40510
R:Gwilliam, R.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.
submitted to the EMBL Data Library, September 1998
A:Reference number: Z21933
A:Accession: T40510
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-736 <GWI>
A:Cross-references: EMBL:AL031534; PIDN:CAA20733.1; GSPDB:GN00067; SPDB:SPBC4F6.13c
A:Experimental source: strain 972h-; cosmid c4F6
C:Genetics:
A:Gene: SPDB:SPBC4F6.13c
A:Map position: 2
C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match 28.0%; Score 46; DB 2; Length 736;
Best Local Similarity 61.1%; Pred. No. 1.6e+02;
Matches 11; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 7 LLCKLPTDSELA PRSXCC 24
II IIIII II I

Db 354 LLPKLTPSELRFPPTRC 371

RESULT 38

MMPTS

laminin beta-2 chain precursor - rat
N:Alternate names: laminin chain B3; S-laminin
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-Dec-1999
C:Accession: S03539
R:Hunter, D.D.; Shah, V.; Merlie, J.P.; Sanes, J.R.
Nature 338, 229-234, 1989
A:Title: A laminin-like adhesive protein concentrated in the synaptic cleft of the neuro
A:Reference number: S03539; MUID:89159410
A:Accession: S03539
A:Molecule type: mRNA
A:Residues: 1-1801 <HUN>
A:Cross-references: EMBL:X16563; NID:g57250; PIDN:CAA34561.1; PID:g57251
A:Complex: Laminins are trimers of an alpha-type, a beta-type, and a gamma-type laminin
C:Function:
A:Description: Interact with cells and with other basement membrane proteins to promote
C:Superfamily: laminin beta-1 chain; laminin-type EGF-like homology
C:Keywords: basement membrane; calcium binding; cell binding; coiled coil; extracellular
F:1-35/Domain: signal sequence #status predicted <SIG>
F:36-1801/Product: laminin beta-2 chain #status predicted <MAT>
F:36-285/Domain: VI <DOM6>
F:286-555/Domain: V <DOM5>
F:286-347/Domain: laminin-type EGF-like homology <LE01>
F:350-410/Domain: laminin-type EGF-like homology <LE02>
F:413-470/Domain: laminin-type EGF-like homology <LE03>
F:473-522/Domain: laminin-type EGF-like homology <LE04>
F:525-555/Domain: laminin-type EGF-like homology #status atypical <LE05>

F:556-784/Domain: IV <DOM4>
F:786-831/Domain: laminin-type EGF-like homology <LE06>
F:788-1196/Domain: III <DOM3>
F:834-877/Domain: laminin-type EGF-like homology <LE07>
F:880-927/Domain: laminin-type EGF-like homology <LE08>
F:930-986/Domain: laminin-type EGF-like homology <LE09>
F:989-1038/Domain: laminin-type EGF-like homology <LE10>
F:1041-1095/Domain: laminin-type EGF-like homology <LE11>
F:1098-1143/Domain: laminin-type EGF-like homology <LE12>
F:1146-1190/Domain: laminin-type EGF-like homology <LE13>
F:1197-1412/Domain: II <DOM2>
F:1197-1412/Region: heptad repeats
F:1413-1445/Domain: alpha <ALP>
F:1446-1801/Region: heptad repeats
F:1446-1801/Domain: I <DOM1>
F:45-50/Disulfide bonds: #status predicted
F:251-371,1088,1252,1311,1351,1502/Binding site: carbohydrate (Asn) (covalent) #status
F:1193,1196,1800/Disulfide bonds: interchain #status predicted

Query Match 27.7%; Score 45.5; DB 1; Length 1801;
Best Local Similarity 40.0%; Pred. No. 3.9e+02;
Matches 10; Conservative 2; Mismatches 10; Indels 3; Gaps 1;

QY 2 RCINQLLCKLPTDSELA PRSXCHC 26
II I III I I I I I

Db 1040 RCTCNL---LGTDPQRCPSDLCCHC 1061

RESULT 39

S24397

stellate protein - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster
C:Date: 19-Feb-1994 #sequence_revision 23-Feb-1996 #text_change 21-Jul-2000
C:Accession: S24397; S08120
R:Divak, K.J.
Genetics 124, 303-316, 1990
A:Title: Detailed structure of the Drosophila melanogaster Stellate genes and their t
A:Reference number: S24397; MUID:90169476
A:Accession: S24397
A:Molecule type: DNA
A:Residues: 1-172 <LIV>
A:Cross-references: EMBL:X15899; NID:g8660; PIDN:CAA33906.1; PID:g295755
C:Genetics:
A:Gene: stellate
A:Cross-references: FlyBase:FBgn0003523
A:Introns: 4/2; 171/2

Query Match 27.4%; Score 45; DB 2; Length 172;
Best Local Similarity 70.0%; Pred. No. 64;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 NQLLCKLPTD 14
II IIIII

Db 21 NQFLCRVPTD 30

RESULT 40

C71415

hypothetical protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)
A:Variety: columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998
C:Accession: C71415
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzensegger, T.; Pohl, T.M.; Terry, N.; G
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomen
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.;
C.; Chludzisz, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis t

Search completed: August 22, 2002, 07:51:24
Job time: 187 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 07:50:21 ; Search time 51.98 Seconds
(without alignments)
59.832 Million cell updates/sec

Title: US-09-856-199-2

Perfect score: 136

Sequence: 1 TGSATPXQCKENLPSLCSXQGLRAENI 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A_Geneseq_032802.*
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	130	95.6	28	21	Human selenoprotein
2	130	95.6	103	21	Human selenoprotein
3	130	95.6	381	21	Human selenoprotein
4	91	66.9	20	21	Human selenoprotein
5	52.5	38.6	566	18	E. coli acetolacta
6	52.5	38.6	604	22	E. coli cellular p
7	52.5	38.2	679	22	Drosophila melanog
8	52	38.2	679	22	Drosophila melanog
9	50.5	37.1	2931	22	Drosophila melanog
10	49	36.0	344	21	Zea mays protein f
11	49	36.0	428	21	Zea mays protein f

12	49	36.0	484	21	AAG43939	Zea mays protein f
13	48.5	35.7	573	22	AAU35664	Haemophilus influe
14	47	34.6	1383	19	AAW48842	Human receptor tyr
15	46	33.8	197	22	ABG00467	Novel human diago
16	46	33.8	895	21	AAG42433	Arabidopsis thalia
17	46	33.8	956	21	AAG42432	Arabidopsis thalia
18	46	33.8	1002	21	AAG42431	Arabidopsis thalia
19	45	33.1	659	21	AAW25107	Eucalyptus grandis
20	44	32.4	65	22	AAO12533	Human polypeptide
21	44	32.4	158	14	AAW32868	17 kD Eimeria tene
22	44	32.4	230	8	AAW70489	Sequence of an ant
23	44	32.4	230	14	AAW32871	Eimeria tenella TA
24	44	32.4	235	6	AAW50014	Sequence of the A4
25	44	32.4	253	8	AAW70487	Sequence of Eimeri
26	44	32.4	253	8	AAW70488	Sequence of the TA
27	44	32.4	253	14	AAW32869	Eimeria tenella pr
28	44	32.4	832	22	ABG32820	Novel human diago
29	44	32.4	932	22	ABG17156	Novel human diago
30	44	32.4	1507	22	ABW38562	Drosophila melanog
31	43	31.6	105	21	AAG56627	Arabidopsis thalia
32	43	31.6	153	21	AAG25073	Arabidopsis thalia
33	43	31.6	177	22	ABW66279	Drosophila melanog
34	43	31.6	208	21	AAG58013	Arabidopsis thalia
35	43	31.6	210	21	AAG55979	Arabidopsis thalia
36	43	31.6	212	21	AAG55978	Arabidopsis thalia
37	43	31.6	326	19	AAW60112	Mycobacterium vacc
38	43	31.6	326	20	AAW14858	M. vaccae antigen
39	43	31.6	344	21	AAW08670	Arabidopsis thalia
40	43	31.6	361	22	AAU57026	Protonibacterium
41	43	31.6	427	21	AAW08669	Arabidopsis thalia
42	43	31.6	443	21	AAW09496	Arabidopsis thalia
43	43	31.6	458	21	AAW08668	Arabidopsis thalia
44	43	31.6	472	21	AAW09495	Arabidopsis thalia
45	43	31.6	481	21	AAW26594	Synechocystis sp p
46	43	31.6	732	22	ABG05396	Novel human diago
47	42.5	31.2	264	18	AAW21778	Protein encoded by
48	42	30.9	136	22	ABG00034	Novel human diago
49	42	30.9	154	22	AAU19789	Human novel extrac
50	42	30.9	207	16	AAW65003	Cattle tissue inhi
51	42	30.9	250	22	AAU19630	Human novel extrac
52	42	30.9	251	17	AAW97978	Stannocalcin alph
53	42	30.9	251	21	AAW55748	Human stannocalci
54	42	30.9	264	22	AAW59018	Drosophila melanog
55	42	30.9	293	20	AAW41255	Adipogenesis inhib
56	42	30.9	293	21	AAW26872	Human adipocytogen
57	42	30.9	293	22	AAW39655	Human protein sequ
58	42	30.9	298	22	AAW33646	Pseudomonas aerugi
59	42	30.9	302	21	AAW67925	Human stannocalci
60	42	30.9	302	22	AAW98971	Human adipogenesis
61	42	30.9	302	22	AAW95330	Human protein sequ
62	42	30.9	302	22	AAW31797	Amino acid sequenc
63	42	30.9	302	22	AAW61621	Human stannocalci
64	42	30.9	372	22	AAW39564	Human polypeptide
65	42	30.9	595	20	AAW27236	Amino acid sequenc
66	42	30.9	761	20	AAW27211	Amino acid sequenc
67	42	30.9	806	22	ABW60866	Drosophila melanog
68	42	30.9	1294	20	AAW39633	Zea mays Rpg1-2 pr
69	42	30.9	1479	18	AAW44118	Murine type C lect
70	42	30.9	1479	18	AAW44119	Human type C lecti
71	42	30.9	1479	21	AAW43151	Human ORFX ORF2915
72	42	30.9	1479	21	AAW23822	Murine E-selectin
73	42	30.9	1479	21	AAW23823	Human type C lecti
74	41.5	30.5	277	21	AAW74968	Neisseria meningit
75	41.5	30.5	332	22	AAW92503	Human protein sequ
76	41.5	30.5	505	22	AAW65678	Novel protein kina
77	41.5	30.5	515	22	AAW60003	Protonibacterium
78	41.5	30.5	658	22	AAE06934	Human membrane-typ
79	41.5	30.5	735	22	AAW95772	Human protein sequ
80	41.5	30.5	756	22	AAW95177	Human protein sequ
81	41.5	30.5	756	22	AAW67427	Amino acid sequenc
82	41.5	30.5	792	22	AAW85782	Human kinase PKIN-
83	41.5	30.5	802	22	AAW41710	Human PRO618 prote
84	41.5	30.5	802	21	AAW44266	Human PRO618 (UNQ3

PA (KAGA) CHEMO-SERO-THERAPEUTIC RES INST.
 XX Hirashima M, Maeda H, Nozaki C;
 XX WPI; 2000-451677/39.
 XX Peptide fragments with cell death inhibitory activity, useful in
 PT preventing and treating apoptosis-associated diseases particularly
 PT caused by stress e.g. Parkinson's disease, Alzheimer's and
 PT arteriosclerosis -
 XX
 XX Disclosure; Page 11; 56pp; Japanese.
 XX
 CC The invention relates to fragments (AAB03183, AAB03184) derived from the
 CC C-terminal 103 amino acids of human selenoprotein P (AAB03188) which act
 CC as inhibitors of apoptosis (programmed cell death), their use in
 CC therapeutic compositions and cell culture media, and antibodies against
 CC them. The invention also relates to a method for studying the phenomenon
 CC of sudden cell death in a human megakaryotic series gemmule cell culture
 CC system involving use of the anti-apoptotic peptides of the invention.
 CC Selenoprotein P is a plasma protein which confers protection against
 CC peroxynitrite-mediated oxidation and nitration. The peptide fragments
 CC can be used as prophylactic and therapeutic agents for
 CC apoptosis-associated diseases, including AIDS (acquired immunodeficiency
 CC syndrome), Parkinson's disease, Alzheimer's disease, arteriosclerosis,
 CC myocardial infarct, cerebral infarct, organ transplant rejection, and
 CC reperfusion injury. They may also be used in the prevention or treatment
 CC of disorders relating to redox reactions, or of disorders of the immune
 CC system. They may additionally be used as an additive for cell culture and
 CC for screening cell death activity in vitro. The present sequence
 CC represents the C-terminal 103 amino acids of human selenoprotein P.
 XX
 XX Sequence 103 AA;
 SQ

Query Match 95.6%; Score 130; DB 21; Length 103;
 Best Local Similarity 92.9%; Pred. No. 2.7e-12;
 Matches 26; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGSATXCKENLPSLCXKQGLRAEENI 28
 ||||| ||||| ||||| ||||| |||||
 Db 34 tgsaitcqckenlpsscscqglraeeni 61

RESULT 3
 AAB03188
 ID AAB03188 standard; peptide; 381 AA.
 XX
 AC AAB03188;
 XX
 DT 23-OCT-2000 (first entry)
 XX
 DE Human selenoprotein P.
 XX
 KW Selenoprotein P; human; programmed cell death inhibitor;
 KW apoptosis inhibitor; antiapoptotic; acquired immunodeficiency syndrome;
 KW AIDS; Parkinson's disease; Alzheimer's disease; myocardial infarct;
 KW cerebral infarct; stroke; reperfusion injury; organ transplant rejection;
 KW arteriosclerosis; redox disorder; immune disorder; cardiovascular;
 KW cerebroprotective; culture medium additive.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..19
 FT /note= "Signal peptide"
 FT Protein 19..381
 FT /note= "Mature human selenoprotein P"
 FT Modified-site 281
 FT /note= "Selenocysteine"
 FT Modified-site 299
 FT /note= "Selenocysteine"
 FT Modified-site 311

FT Modified-site /note= "Selenocysteine"
 326
 FT Modified-site /note= "Selenocysteine"
 333
 FT Modified-site /note= "Selenocysteine"
 348
 FT Modified-site /note= "Selenocysteine"
 350
 FT Modified-site /note= "Selenocysteine"
 357
 FT Modified-site /note= "Selenocysteine"
 359
 FT Modified-site /note= "Selenocysteine"
 359
 XX
 XX WO2000031131-A1.
 PN
 XX 02-JUN-2000.
 PD
 XX 12-NOV-1999; 99WO-JP06322.
 PF
 XX 19-NOV-1998; 98JP-0347863.
 PR
 XX (KAGA) CHEMO-SERO-THERAPEUTIC RES INST.
 PA
 XX Hirashima M, Maeda H, Nozaki C;
 PI
 XX WPI; 2000-451677/39.
 DR
 XX Peptide fragments with cell death inhibitory activity, useful in
 PT preventing and treating apoptosis-associated diseases particularly
 PT caused by stress e.g. Parkinson's disease, Alzheimer's and
 PT arteriosclerosis -
 XX
 XX Disclosure; Page 11; 56pp; Japanese.
 XX
 CC The invention relates to fragments (AAB03183, AAB03184) derived from the
 CC C-terminal 103 amino acids of human selenoprotein P (AAB03188) which act
 CC as inhibitors of apoptosis (programmed cell death), their use in
 CC therapeutic compositions and cell culture media, and antibodies against
 CC them. The invention also relates to a method for studying the phenomenon
 CC of sudden cell death in a human megakaryotic series gemmule cell culture
 CC system involving use of the anti-apoptotic peptides of the invention.
 CC Selenoprotein P is a plasma protein which confers protection against
 CC peroxynitrite-mediated oxidation and nitration. The peptide fragments
 CC can be used as prophylactic and therapeutic agents for
 CC apoptosis-associated diseases, including AIDS (acquired immunodeficiency
 CC syndrome), Parkinson's disease, Alzheimer's disease, arteriosclerosis,
 CC myocardial infarct, cerebral infarct, organ transplant rejection, and
 CC reperfusion injury. They may also be used in the prevention or treatment
 CC of disorders relating to redox reactions, or of disorders of the immune
 CC system. They may additionally be used as an additive for cell culture and
 CC for screening cell death activity in vitro. The present sequence
 CC represents the C-terminal 103 amino acids of human selenoprotein P.
 XX
 XX Sequence 381 AA;
 SQ

The invention relates to fragments (AAB03183, AAB03184) derived from the
 C-terminal 103 amino acids of human selenoprotein P (AAB03188) which act
 as inhibitors of apoptosis (programmed cell death), their use in
 therapeutic compositions and cell culture media, and antibodies against
 them. The invention also relates to a method for studying the phenomenon
 of sudden cell death in a human megakaryotic series gemmule cell culture
 system involving use of the anti-apoptotic peptides of the invention.
 Selenoprotein P is a plasma protein which confers protection against
 peroxynitrite-mediated oxidation and nitration. The peptide fragments
 can be used as prophylactic and therapeutic agents for
 apoptosis-associated diseases, including AIDS (acquired immunodeficiency
 syndrome), Parkinson's disease, Alzheimer's disease, arteriosclerosis,
 myocardial infarct, cerebral infarct, organ transplant rejection, and
 reperfusion injury. They may also be used in the prevention or treatment
 of disorders relating to redox reactions, or of disorders of the immune
 system. They may additionally be used as an additive for cell culture and
 for screening cell death activity in vitro. The present sequence
 represents human selenoprotein P.

Query Match 95.6%; Score 130; DB 21; Length 381;
 Best Local Similarity 92.9%; Pred. No. 1.2e-11;
 Matches 26; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGSATXCKENLPSLCXKQGLRAEENI 28
 ||||| ||||| ||||| ||||| |||||
 Db 312 tgsaitcqckenlpsscscqglraeeni 339

RESULT 4
 AAB03189
 ID AAB03189 standard; peptide; 20 AA.
 XX
 AC AAB03189;
 XX
 DT 23-OCT-2000 (first entry)

XX DE Human selenoprotein P C-terminal fragment, SEQ ID NO:7.
 XX KW Selenoprotein P fragment; human; programmed cell death inhibitor;
 KW apoptosis inhibitor; antiapoptotic; acquired immunodeficiency syndrome;
 KW AIDS; Parkinson's disease; Alzheimer's disease; myocardial infarct;
 KW cerebral infarct; stroke; reperfusion injury; organ transplant rejection;
 KW arteriosclerosis; redox disorder; immune disorder; cardiovascular;
 KW cerebroprotective; culture medium additive.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT Modified-site 7
 FT Modified-site 19 /note= "Selenocysteine"
 FT Modified-site 19 /note= "Selenocysteine"
 XX WO2000031131-A1.
 XX 02-JUN-2000.
 XX 12-NOV-1999; 99WO-JP06332.
 XX 19-NOV-1998; 98JP-0347863.
 XX (KAGA) CEMO-SERO-THERAPEUTIC RES INST.
 XX Hirashima M, Maeda H, Nozaki C;
 XX WPI; 2000-451677/39.
 XX Peptide fragments with cell death inhibitory activity, useful in
 PT preventing and treating apoptosis-associated diseases particularly
 PT caused by stress e.g. Parkinson's disease, Alzheimer's and
 PT arteriosclerosis -
 XX Example 7; Page 25; 56pp; Japanese.
 XX The invention relates to fragments (AAB03183, AAB03184) derived from the
 CC C-terminal 103 amino acids of human selenoprotein P (AAB03188) which act
 CC as inhibitors of apoptosis (programmed cell death), their use in
 CC therapeutic compositions and cell culture media, and antibodies against
 CC them. The invention also relates to a method for studying the phenomenon
 CC of sudden cell death in a human megakaryotic series gemmule cell culture
 CC system involving use of the anti-apoptotic peptides of the invention.
 CC Selenoprotein P is a plasma protein which confers protection against
 CC peroxynitrite-mediated oxidation and nitration. The peptide fragments
 CC can be used as prophylactic and therapeutic agents for
 CC apoptosis-associated diseases, including AIDS (acquired immunodeficiency
 CC syndrome), Parkinson's disease, Alzheimer's disease, arteriosclerosis,
 CC myocardial infarct, cerebral infarct, organ transplant rejection, and
 CC reperfusion injury. They may also be used in the prevention or treatment
 CC of disorders relating to redox reactions, or of disorders of the immune
 CC system. They may additionally be used as an additive for cell culture and
 CC for screening cell death activity in vitro. The present sequence
 CC represents a C-terminal fragment of human selenoprotein P used in an
 CC exemplification of the invention.
 XX Sequence 20 AA;

Query Match 66.9%; Score 91; DB 21; Length 20;
 Best Local Similarity 90.0%; Pred. No. 3.7e-07;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TGSALTQCKENLPSLCSXQ 20
 ||||| ||||| ||||| ||||| |||||
 Db 1 tgsaltcqckenlpslcsq 20

RESULT 5
 AAW22461

ID AAW22461 standard; Protein; 566 AA.
 XX AC AAW22461;
 XX DT 25-SEP-1997 (first entry)
 XX DE E. coli acetolactate synthase isozyme III (sub-sequence G).
 XX KW Acetolactate synthase; ALS; herbicide resistance; transgenic plant;
 KW crop protection.
 XX OS Escherichia coli.
 XX PN US5605011-A.
 XX PD 25-FEB-1997.
 XX PF 26-AUG-1986; 86US-0900609.
 XX PR 27-JUL-1987; 87IL-0083348.
 XX PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX Bedbrook JR, Chaleff RS, Falco SC, Mazur BJ, Somerville CR;
 PT Yadav NS;
 XX WPI; 1997-153232/14.
 XX Use of mutant acetolactate synthase genes - for transforming plants
 PT for resistance to sulphonylurea, triazopyrimidine sulphonamide
 PT and imidazolinone herbicides.
 XX Claim 1; Fig 6a-f; 63pp; English.
 CC Amino acid sub-sequences A-G respectively correspond to the large
 CC subunits of acetolactate synthase (ALS) isozymes I, II and III
 CC (AAW22459-61) from E. coli, and wild-type ALS proteins of yeast
 CC (AAW22458), Arabidopsis thaliana (AAW22457) and tobacco (AAW22455 and
 CC AAW22456). Comparison of these substantially conserved sequences
 CC with those of herbicide-resistant ALS enzymes (see also AAW22462-64)
 CC reveals the location of amino acid substituents that lead to herbicide
 CC resistance. In sub-sequence G, such a substituent is located at the
 CC sigma-1 position (any amino acid other than methionine). A nucleic
 CC acid encoding an ALS from any source can be mutated so that the
 CC encoded enzyme contains this amino acid substituent. Transformation of
 CC herbicide sensitive plants or plant cells with the nucleic acid
 CC results in resistance to the herbicide.
 XX Sequence 566 AA;
 SQ Query Match 38.6%; Score 52.5; DB 18; Length 566;
 Best Local Similarity 40.6%; Pred. No. 14;
 Matches 13; Conservative 2; Mismatches 8; Indels 9; Gaps 1;
 Qy 2 GSAITXQCKE-----NLPSLCSXQGLRA 24
 ||||| : ||||| : ||||| : ||||| :
 Db 215 ggaitagchqqlketvealnlpvcsimglga 246
 RESULT 6
 AAU34429
 ID AAU34429 standard; Protein; 604 AA.
 XX AC AAU34429;
 XX DT 14-FEB-2002 (first entry)
 XX DE E. coli cellular proliferation protein #10.
 XX KW Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.
 XX

```

OS Escherichia coli.
XX WO200170955-A2.
XX
XX PD 27-SEP-2001.
XX
XX PF 21-MAR-2001; 2001WO-US09180.
XX
XX PR 21-MAR-2000; 2000US-191078P.
XX PR 23-MAY-2000; 2000US-206848P.
XX PR 26-MAY-2000; 2000US-207727P.
XX PR 23-OCT-2000; 2000US-242578P.
XX PR 27-NOV-2000; 2000US-253625P.
XX PR 22-DEC-2000; 2000US-257931P.
XX PR 16-FEB-2001; 2001US-269308P.
XX
XX PA (ELIT-) ELITRA PHARM INC.
XX
XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX PI Yamamoto RT, Xu HH;
XX
XX DR WPI; 2001-611495/70.
XX DR N-PSDB; AAS52288.
XX
XX PT New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX PS Example 3; Seq ID No 10022; 511pp; English.
XX
XX CC The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the
XX CC genes, their use in the discovery of novel antibiotics, the essential
XX CC genes themselves and the encoded proteins. The prokaryotes used are
XX CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX CC invention is also useful for the identification of potential new targets
XX CC for antibiotic development. The antisense nucleic acids can also be used
XX CC to identify proteins used in proliferation, to express these proteins,
XX CC and to obtain antibodies capable of binding to the expressed proteins.
XX CC The proteins can be used to screen compounds in rational drug discovery
XX CC programmes. The antisense nucleic acid sequence is also useful to screen
XX CC for homologous nucleic acids which are required for cell proliferation in
XX CC a wide variety of organisms. The present sequence represents an
XX CC essential prokaryotic cellular proliferation protein.
XX CC Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 604 AA;

Query Match 38.6%; Score 52.5; DB 22; Length 604;
Best Local Similarity 40.6%; Pred. No. 15;
Matches 13; Conservative 2; Mismatches 8; Indels 9; Gaps 1;

QY 2 GSATYXQCKE-----NLPSCSXQGLRA 24
| | | | | : | | | | |
Db 245 ggaitagchqqlketvealnlpvvcslmg_lga 276

RESULT 7
ABBG2504
ID ABB62504 standard; Protein; 670 AA.
XX
XX AC ABB62504;
XX
XX DT 26-MAR-2002 (first entry)
XX
XX DE Drosophila melanogaster polypeptide SEQ ID NO 14304.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.

```

```

XX Drosophila melanogaster.
XX OS WO200171042-A2.
XX PN
XX PD 27-SEP-2001.
XX
XX PF 23-MAR-2001; 2001WO-US09231.
XX
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX
XX PA (PEKE ) PE CORP NY.
XX
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX PI WPI; 2001-656860/75.
XX DR N-PSDB; ABL06607.
XX
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions -
XX
XX PS Disclosure; SEQ ID NO 14304; 21pp + Sequence Listing; English.
XX
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins
XX CC (ABB57737-ABB72072).
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 670 AA;

Query Match 38.2%; Score 52; DB 22; Length 670;
Best Local Similarity 61.5%; Pred. No. 20;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 ITXQCKENLPISLC 17
: : | | | | : | | |
Db 148 vharckenvpislc 160

RESULT 8
ABBG7010
ID ABB67010 standard; Protein; 679 AA.
XX
XX AC ABB67010;
XX
XX DT 26-MAR-2002 (first entry)
XX
XX DE Drosophila melanogaster polypeptide SEQ ID NO 27822.
XX
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX
XX OS Drosophila melanogaster.
XX
XX PN WO200171042-A2.
XX
XX PD 27-SEP-2001.
XX
XX PF 23-MAR-2001; 2001WO-US09231.
XX
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX

```

PA (PEKE) PE CORP NY.
PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
DR N-PSDB; ABL11113.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 27822; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 679 AA;

Query Match 38.2%; Score 52; DB 22; Length 679;
Best Local Similarity 61.5%; Pred. No. 21;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 5 ITXOCKENLPSLC 17
Db 157 vharckenvpslc 169
:
:|||||:
:
:

RESULT 9
ABB68229
ID ABB68229 standard; Protein: 2931 AA.
XX
AC ABB68229;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 31479.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL12332.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 31479; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2931 AA;

Query Match 37.1%; Score 50.5; DB 22; Length 2931;
Best Local Similarity 27.5%; Pred. No. 1.9e+02;
Matches 11; Conservative 7; Mismatches 9; Indels 13; Gaps 1;

Qy 2 GSAITXOCKENLPSLC SXQG-----LRAEENI 28
Db 57 gavierngcdnlasvctesgdcncycrshgcnfinsqenl 96
:
:|||||:
:
:

RESULT 10
AAG43941
ID AAG43941 standard; Protein: 344 AA.
XX
AC AAG43941;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 54981.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 05-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 30-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.

```
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0162142.

Query Match 36.08; Score 49; DB 21; Length 344;
Best Local Similarity 40.08; Pred. No. 27;
Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 9 CKENPLSCSXQGLRAEENI 28
   I: I I I I I I I I I I
Db 226 cecnfpsichsgmtqeel 245

RESULT 11
AAG43940
ID AAG43940 standard; Protein; 428 AA.
XX AC AAG43940;
XX DT 18-OCT-2000 (first entry)
XX DE
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence; corn.
XX OS
XX ZE zeae mayssubsp. mayss.
XX PN EPI033405-A2.
XX PD
XX PF 06-SEP-2000.
XX FF
XX FF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 23-APR-1999; 99US-0130891.
XX PR 28-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 04-MAY-1999; 99US-0132407.
XX PR 04-MAY-1999; 99US-0132484.
XX PR 05-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 06-MAY-1999; 99US-0132487.
XX PR 07-MAY-1999; 99US-0132863.
XX PR 11-MAY-1999; 99US-0134256.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 14-MAY-1999; 99US-0134370.
XX PR 18-MAY-1999; 99US-01343768.
XX PR 19-MAY-1999; 99US-0134941.
XX PR 20-MAY-1999; 99US-0135124.
XX PR 21-MAY-1999; 99US-0135353.
XX PR 24-MAY-1999; 99US-0135629.
XX PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
```

PR	04-AUG-1999;	99US-0147204.	PR	29-OCT-1999;	99US-0162142.
PR	04-AUG-1999;	99US-0147302.		Query Match	36.0%; Score 49; DB 21; Length 428;
PR	05-AUG-1999;	99US-0147192.		Best Local Similarity	40.0%; Pred. No. 35;
PR	05-AUG-1999;	99US-0147260.		Matches	8; Conservative 4; Mismatches 8; Indels 0; Gaps
PR	06-AUG-1999;	99US-0147303.			
PR	06-AUG-1999;	99US-0147416.			
PR	08-AUG-1999;	99US-0147493.			
PR	08-AUG-1999;	99US-0147935.			
PR	10-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			
PR	12-AUG-1999;	99US-0148341.			
PR	13-AUG-1999;	99US-0148565.			
PR	13-AUG-1999;	99US-0148684.			
PR	16-AUG-1999;	99US-0149368.			
PR	17-AUG-1999;	99US-0149175.			
PR	18-AUG-1999;	99US-0149426.			
PR	20-AUG-1999;	99US-0149722.			
PR	20-AUG-1999;	99US-0149723.			
PR	20-AUG-1999;	99US-0149929.			
PR	23-AUG-1999;	99US-0149902.			
PR	23-AUG-1999;	99US-0149930.			
PR	25-AUG-1999;	99US-0150566.			
PR	26-AUG-1999;	99US-0150884.			
PR	27-AUG-1999;	99US-0151065.			
PR	27-AUG-1999;	99US-0151066.			
PR	27-AUG-1999;	99US-0151080.			
PR	30-AUG-1999;	99US-0151303.			
PR	31-AUG-1999;	99US-0151438.			
PR	01-SEP-1999;	99US-0151930.			
PR	07-SEP-1999;	99US-0152363.			
PR	10-SEP-1999;	99US-0153070.			
PR	13-SEP-1999;	99US-0153758.			
PR	15-SEP-1999;	99US-0154018.			
PR	16-SEP-1999;	99US-0154039.			
PR	20-SEP-1999;	99US-0154779.			
PR	22-SEP-1999;	99US-0155139.			
PR	23-SEP-1999;	99US-0155486.			
PR	24-SEP-1999;	99US-0155659.			
PR	28-SEP-1999;	99US-0156458.			
PR	29-SEP-1999;	99US-0156596.			
PR	04-OCT-1999;	99US-0157117.			
PR	05-OCT-1999;	99US-0157753.			
PR	08-OCT-1999;	99US-0157865.			
PR	07-OCT-1999;	99US-0158029.			
PR	08-OCT-1999;	99US-0158232.			
PR	12-OCT-1999;	99US-0158369.			
PR	13-OCT-1999;	99US-0159293.			
PR	13-OCT-1999;	99US-0159294.			
PR	13-OCT-1999;	99US-0159295.			
PR	14-OCT-1999;	99US-0159329.			
PR	14-OCT-1999;	99US-0159330.			
PR	14-OCT-1999;	99US-0159331.			
PR	14-OCT-1999;	99US-0159637.			
PR	14-OCT-1999;	99US-0159638.			
PR	18-OCT-1999;	99US-0159584.			
PR	21-OCT-1999;	99US-0160741.			
PR	21-OCT-1999;	99US-0160767.			
PR	21-OCT-1999;	99US-0160768.			
PR	21-OCT-1999;	99US-0160770.			
PR	21-OCT-1999;	99US-0160814.			
PR	21-OCT-1999;	99US-0160815.			
PR	22-OCT-1999;	99US-0160980.			
PR	22-OCT-1999;	99US-0160981.			
PR	22-OCT-1999;	99US-0160989.			
PR	25-OCT-1999;	99US-0161404.			
PR	25-OCT-1999;	99US-0161405.			
PR	25-OCT-1999;	99US-0161406.			
PR	26-OCT-1999;	99US-0161359.			
PR	26-OCT-1999;	99US-0161360.			
PR	26-OCT-1999;	99US-0161361.			
PR	28-OCT-1999;	99US-0161920.			
PR	28-OCT-1999;	99US-0161992.			
PR	28-OCT-1999;	99US-0161993.			
PR	04-AUG-1999;	99US-0147204.	PR	29-OCT-1999;	99US-0162142.
PR	04-AUG-1999;	99US-0147302.		Query Match	36.0%; Score 49; DB 21; Length 428;
PR	05-AUG-1999;	99US-0147192.		Best Local Similarity	40.0%;

PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 36.0%; Score 49; DB 21; Length 484;
Best Local Similarity 40.0%; Pred. No. 40;
Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 9 CKENLPSCSXQGLRAENI 28

XX PT New isolated receptor tyrosine kinase genes - which are expressed in
 PT neuronal tissues and tumour cells, useful as targets for
 PT neurodegenerative disorders or cancers
 XX Example 1: Fig 7A: 87pp; English.
 XX This is the amino acid sequence of human LMR1-h, deduced from a
 CC LMR1-h cDNA clone (see AAV32449), with signal peptide sequences
 CC supplied from rat LMR1-r. Novel rat, human and mouse LMR1,
 CC LMR2 and LMR3 (see AAW4841-49) define a novel family of receptors
 CC that are structurally related to receptor tyrosine kinases (RTKs).
 CC They all share the distinct motifs that typically characterise this
 CC class of enzymes, but possess extremely short extracellular domains
 CC and have C-terminal tails of unprecendented length among RTKs.
 CC Expression of LMR1 and LMR3 is highly restricted to neuronal
 CC tissues with minimal expression in other adult or embryonic organs
 CC or in human tumour cell lines. LMR2 expression is limited to adult
 CC neuronal tissues, but is also very abundantly expressed in other
 CC non-neuronal foetal tissues and in numerous tumour cell lines.
 CC Based on restricted expression of all 3 LMRs to adult neuronal
 CC tissues and the up regulation of LMR2 in a wide variety of tumour
 CC cell lines, these proteins may be critical targets for
 CC neurodegenerative disorders or cancer. Anti-LMR antibodies may be
 CC used for detecting neurodegenerative diseases or cancer
 XX SQ Sequence 1383 AA;
 Query Match 34.6%; Score 47; DB 19; Length 1383;
 Best Local Similarity 64.3%; Pred. No. 2.7e+02;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 9 KRENPLSLCSXOGL 22
 Db 750 ccpgphicsaql 763
 RESULT 15
 ABG00467
 ID ABG00467 standard; Protein; 197 AA.
 XX AC ABG00467;
 XX DT 13-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #458.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX OS Homo sapiens.
 XX PN W0200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX DR WPI; 2001-639362/73.
 DR N-PSDB; AAS64654.
 XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX Claim 20; SEQ ID No 30826; 103pp; English.
 PS The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 197 AA;
 Query Match 33.8%; Score 46; DB 22; Length 197;
 Best Local Similarity 47.4%; Pred. No. 40;
 Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 Qy 10 KENPLSLCSXOGLRAEENT 28
 Db 156 eenhpsilrtqglkaadgi 174
 RESULT 16
 AAG42433
 ID AAG42433 standard; Protein; 895 AA.
 XX AC AAG42433;
 XX DT 18-OCT-2000 (first entry)
 XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 52919.
 XX KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX OS Arabidopsis thaliana.
 XX PN EP1033405-A2.
 XX PD 06-SEP-2000.
 XX PF 25-FEB-2000; 2000EP-0301439.
 XX PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.


```
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151085.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.

PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 33.8%; Score 46; DB 21; Length 956;
Best Local Similarity 60.0%; Pred. No. 2.5e+02;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TGSATYXCKENLPS 15
Db 391 sgsmiddgcseklps 405
:| | | | | | | | | |
| | | | | | | | | |

RESULT 18
AAG42431
ID AAG42431 standard; Protein; 1002 AA.
XX
AC AAG42431;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 52917.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
```

PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139432.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152353.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.

PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 33.8%; Score 46; DB 21; Length 1002;
Best Local Similarity 60.0%; Pred. No. 2.7e+02;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TGSATXOCKENLPS 15
Db :||| ||| |||
437 sgsmiddqcseklps 451

RESULT 19
AAB25107
ID AAB25107 standard; Protein; 659 AA.
XX
AC AAB25107;
XX
DT 27-NOV-2000 (first entry)
XX
DE Eucalyptus grandis cell signalling involved protein SEQ ID NO:75.
XX
KW Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
KW plant cell signalling; modulation; transgenic plant; pathogen; growth;
KW environmental change; development; cell proliferation; differentiation;
KW elongation; survival; disease resistance; nutrient metabolism.
XX
OS Eucalyptus grandis.
XX
PN WO200042171-A1.
XX
PD 20-JUL-2000.
XX
PF 11-JAN-2000; 2000WO-US00724.
XX
PR 12-JAN-1999; 99US-0228986.
PR 01-NOV-1999; 99US-0162866.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI Strabala TJ, Nieuwenhuizen NJ;
XX
DR WPI; 2000-476052/41.
XX

Isolated polynucleotide encoding a polypeptide involved in cell signaling used for generating transgenic plants with modified responses to external signals -

Claim 3; Page 71-73; 527pp; English.

CC AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide and protein sequences isolated from eucalyptus (*Eucalyptus grandis*) or pine (*Pinus radiata*) also known as Monterey pine). The protein sequences are involved in cell signalling. The polynucleotide and protein sequences can be used to modify the response of plant cells to external signals e.g. environmental changes or pathogens during the growth and development of a plant. They can be used to modify cell proliferation, differentiation, elongation and survival, resistance to disease and nutrient metabolism. Examples of modifications which can be produced are altered fruit ripening and senescence of leaves and flowers e.g. to delay senescence and prolong the life of cut flowers or enhance senescence of reproductive organs to engineer sterile plants. Other modifications can be used to delay senescence in selected cell types or organs providing fruit and vegetables which have a longer shelf life between harvest and consumption, or to decrease branching frequency in forest tree species giving long stretches of valuable knot-free clear wood which can be used in solid timber furniture and veneers.

XX Sequence 659 AA;

Query Match 33.1%; Score 45; DB 21; Length 659;
Best Local Similarity 56.2%; Pred. No. 2.3e+02;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 12 NLPSCSXQGLRAEEN 27
Db ||||| ||| |::| |
49 nlpspcswagvcqegn 64

RESULT 20
AAO12533
ID AAO12533 standard; Protein; 65 AA.
XX
AC AAO12533;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 26425.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR N-PSDB; AAI92464.
XX

Isolated nucleic acids and polypeptides, useful for preventing diagnosing and treating e.g. leukaemia, inflammation and immune disorders -

Claim 20; SEQ ID NO 26425; 1399pp + Sequence Listing; English.

CC The invention relates to human polynucleotides (AAI79941-AAI93841) and the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation.
CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 65 AA;

Query Match 32.4%; Score 44; DB 22; Length 65;
Best Local Similarity 57.1%; Pred. No. 22;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 12 NLPSCSXQGLRAE 25

Db 1 nrpelcdyqgired 14
| | | | | | | | | |

RESULT 21

AAR32868
ID AAR32868 standard; protein; 158 AA.

XX AC AAR32868;

DT 11-JUN-1993 (first entry)

XX 17 kD Eimeria tenella TA4 antigen component.

XX Coccidiosis; chicken; poultry; Apicomplexa; vaccine.

XX Eimeria tenella.

Key	Location/Qualifiers
Peptide	2..14
Peptide	/note= "determined directly"
Peptide	10..35
Peptide	/note= "chymotryptic fragment CH3"
Peptide	14..39
Peptide	/note= "V8 protease fragment V8"
Peptide	16..48
Peptide	/note= "Cyanogen bromide fragment CNI"
Peptide	36..53
Peptide	/note= "chymotryptic fragment CH2"
Peptide	45..52
Peptide	/note= "V8 protease fragment V6"
Peptide	47..70
Peptide	/note= "Endoprotease Arg-C fragment R2"
Peptide	3..74
Peptide	/note= "V8 protease fragment V4"
Peptide	53..74
Peptide	/note= "V8 protease fragment V2"
Peptide	65..81
Peptide	/note= "chymotryptic fragment CH5"
Peptide	75..93
Peptide	/note= "V8 protease fragment V5"
Peptide	75..111
Peptide	/note= "V8 protease fragment V1"
Peptide	69..100
Peptide	/note= "endoprotease Arg-C fragment R1"
Peptide	102..121
Peptide	/note= "chymotryptic fragment CH2"
Peptide	103..130
Peptide	/note= "endoprotease Arg-C R4"
Peptide	120..137
Peptide	/note= "endoprotease Arg-C R2"
Peptide	126..132
Peptide	/note= "chymotryptic fragment CH4"
Peptide	133..158
Peptide	/note= "chymotryptic fragment CH1"

US5187080-A.

16-FEB-1993.

XX 05-JUN-1984; 84US-0617483.

XX 05-JUN-1984; 84US-0617483.

XX 16-MAY-1985; 85US-0734085.

XX 06-DEC-1985; 85US-0805824.

XX 24-NOV-1987; 87US-0125012.

XX 21-JAN-1992; 92US-0824700.

(SOLV) SOLVAY & CIE.

XX Andrews WH, Brothers VM, Files JG, Gore TC, Kuhn I;

PI McCaman MT, Newman K2, Paul LS, Sias SR, Tedesco JL;

XX

DR WPI; 1993-075726/09.

XX Nucleic acid encoding Eimeria tenella proteins - capable of
PT inducing immune response in chickens, useful in vaccine

XX Claim 1; Fig 1; 62pp; English.

XX The 17,000 Dalton peptide component of the TA4 antigen was purified
CC from Eimeria tenella sporulated oocysts. Fractions containing the
CC 17 kD polypeptide also contained a peptide of 8 kD. In sporocyst
CC membrane preparations the 8,000 Dalton protein appears to be linked
CC to the 17,000 Dalton protein by a disulphide bond. The N-terminus
CC of the 17 kD protein was blocked so the protein was reduced and
CC alkylated and then digested with CNBr, V8 protease, chymotrypsin
CC and endoprotease Arg-C. These peptide subfragments were sequenced
CC and the amino acid sequence of the 17kD protein was determined from
CC overlapping fragments. The protein induces an immune response in
CC chickens and can be used in the development of vaccines to protect
CC poultry against E.tenella.

SQ Sequence 158 AA;

Query Match 32.4%; Score 44; DB 14; Length 158;

Best Local Similarity 40.0%; Pred. No. 63;

Matches 10; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 3 SAITXQCKENLPISLCSXQGLRAEEN 27

Db 5 tavltdckeamnkirkagipafed 29

RESULT 22

AAP70489

ID AAP70489 standard; Protein; 230 AA.

XX AC AAP70489;

XX 02-APR-1991 (first entry)

XX Sequence of an antigenic protein derived from Eimeria tenella

DE encoded by pTCD26.

XX Poultry vaccine; Eimeria tenella vaccine; Eimeria necatrix vaccine.

XX Eimeria tenella.

XX AU8665869-A.

XX 04-JUN-1987.

XX 01-DEC-1986; 86AU-0065869.

XX 11-DEC-1985; 85US-0808013.

XX 03-DEC-1985; 85US-0805301.

XX 06-DEC-1985; 85US-0805824.

XX 11-DEC-1985; 85US-0807497.

(SOLV) SOLVAY & CIE.

XX Newman K2, Gore TC, Tedesco JL, Peterson GR, Brothers VM;

PI Paul LS, Chang RJ, Andrews WH;

PI Kuhn I, McCaman MT, Files JG, Sias SR;

XX WPI; 1987-199027/29.

XX N-PSDB; AAN70806.

XX New polypeptide for protecting chickens against coccidiosis - are

PT obtd. by recombinant DNA procedures involving total genomic DNA

XX from eimeria maxima oocysts.

XX Claim 4; Fig 7; 255pp; English.

XX

CC Eimeria tenella polypeptides are used to immunise chickens against
 CC infection against Eimeria strains, so that coccidiosis is prevented
 CC They are used in vaccines, eg in a dose to provide over 20
 CC microgram/kg. The monoclonal antibody or its anti-idiotypic antibody
 CC may also be administered to confer protection. The sequences in Fig
 CC 3 (AAN70804, AAR70487) are very similar (identical apart from DNA
 CC ambiguities?) to the sequences in Fig 5 (AAN70805, AAP70488).

XX
 SQ Sequence 230 AA;

Query Match 32.4%; Score 44; DB 8; Length 230;
 Best Local Similarity 40.0%; Pred. No. 98;
 Matches 10; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Oy 3 SAITXQCKENLPSCSXQGLRAEEN 27
 :|:| ||| : | ||| :
 Db 5 tavltdckeamnkirkaglpafed 29

RESULT 23
 AAR32871
 ID AAR32871 standard; Protein; 230 AA.

XX
 AC AAR32871;

XX
 DT 11-JUN-1993 (first entry)

XX
 DE Eimeria tenella TA4 pre-pro-antigen gene.

XX
 KW Coccidiosis; chicken; poultry; Apicomplexa; vaccine.

XX
 OS Eimeria tenella.

XX
 FH Key Location/Qualifiers

FT Protein 1..158 /note= "17 kD component"

FT Protein 162..230 /note= "8 kD component"

FT
 US187080-A.

XX
 PD 16-FEB-1993.

XX
 PF 05-JUN-1984; 84US-0617483.

XX
 PR 05-JUN-1984; 84US-0617483.

XX
 PR 16-MAY-1985; 85US-0734085.

XX
 PR 06-DEC-1985; 85US-0805824.

XX
 PR 24-NOV-1987; 87US-0125012.

XX
 PR 21-JAN-1992; 92US-0824700.

XX
 PA (SOLV) SOLVAY & CIE.

XX
 PI Andrews WH, Brothers VM, Files JG, Gore TC, Kuhn I;

PI McCaman MT, Newman KZ, Paul LS, Sias SR, Tedesco JL;

XX
 DR WPI; 1993-075726/09.

XX
 DR N-PSDB; RAQ37483.

XX
 PT Nucleic acid encoding Eimeria tenella proteins - capable of

PT inducing immune response in chickens, useful in vaccine

XX
 PS Example 9; Fig 7; 62pp; English.

XX
 CC A cDNA library was prepared from mRNA isolated from sporulated E.
 CC tenella oocysts. The library was screened with the genomic clone
 CC of the TA4 gene sequence. One positive clone, designated pTCD26,
 CC was sequenced and the DNA sequence agreed with that predicted from
 CC the genomic clone. The predicted amino acid sequence from the cDNA
 CC also agreed with the TA4 antigen amino acid sequence obtained by
 CC protein microsequencing. Clone pTCD26 was transformed into E.coli
 CC JM83 and the transformant strain was deposited as JM83/pTCD26

CC (ATCC 53315). The two peptides of the TA4 antigen (17 kD and 8 kD)
 CC are apparently encoded by a contiguous nucleotide sequence, and at
 CC least one proteolytic step occurs to generate the 8 kD peptide by
 CC removal of a tripeptide. Both subunits are capable of inducing an
 CC immune response in chickens to confer protection against E.tenella.

XX
 SQ Sequence 230 AA;

Query Match 32.4%; Score 44; DB 14; Length 230;
 Best Local Similarity 40.0%; Pred. No. 98;
 Matches 10; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Oy 3 SAITXQCKENLPSCSXQGLRAEEN 27
 :|:| ||| : | ||| :
 Db 5 tavltdckeamnkirkaglpafed 29

RESULT 24
 AAP50014
 ID AAP50014 standard; Protein; 235 AA.

XX
 AC AAP50014;

XX
 DT 04-SEP-1991 (first entry)

XX
 DE Sequence of the A4 antigen including the signal peptide, 17000
 DE dalton and 8000 dalton polypeptide components.

XX
 KW Coccidiosis; poultry; chicken; antigen; vaccine.

XX
 OS Eimeria tenella.

XX
 FH Key Location/Qualifiers

FT Peptide 1..23 /label= signal

FT Peptide 24..181 /label= 17,000 dalton

FT Peptide 182..235 /label= 8,000 dalton

XX
 PN EP164176-A.

XX
 PD 11-DEC-1985.

XX
 PF 05-JUN-1985; 85EP-0122355.

XX
 PR 16-MAY-1985; 85US-0734085.

XX
 PR 05-JUN-1984; 84US-0617483.

XX
 PR 06-DEC-1985; 85US-0805824.

XX
 PA (SOLV) SOLVAY & CIE.

XX
 PI Newman KZ, Tedesco JL, Simonson RR, Gore TC, Petersen GR;

PI Brothers VM, Files JG, Paul LS;

XX
 DR WPI; 1985-312086/50.

XX
 DR N-PSDB; AAN50014.

XX
 PT New antigenic protein derivs. from Eimeria species - useful as

PT vaccines and for raising specific monoclonal antibodies

XX
 PS Claim 1; Fig 3; 83pp; English.

XX
 CC The inventors claim a new purified antigenic protein (see AAP50014),
 CC which has mol. wt. about 25000 and consists of two polypeptides
 CC joined by a disulphide bond. One polypeptide has mol. wt about
 CC 17000 and has a blocked N-terminal AA; the other has mol. wt. about
 CC 8000. Also claimed are nucleic acids coding for the antigenic
 CC protein and its fragments. The antigenic protein confers immunity
 CC in chickens against infection by Eimeria tenella or E. necatrix, so
 CC is useful as a vaccine. The pref. dose is 0.1 microg to 1 mg. used
 CC with a standard carrier.

```

XX SQ Sequence 235 AA;
Query Match 32.4%; Score 44; DB 6; Length 235;
Best Local Similarity 40.0%; Pred. No. 1.e+02;
Matches 10; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 3 SAITXQCKENLPISLCSXQGLRAEEN 27
Db 28 tavltdckeamnkirkgaaglpafed 52
:|:| ||| : | || | | :

RESULT 25
AAP70487
ID AAP70487 standard; protein; 253 AA.
XX
AC AAP70487;
XX
DT 02-APR-1991 (first entry)
XX
DE Sequence of Eimeria tenella T4A antigen encoded by the BglII/EcoRI
DE DNA fragment of genomic clone 108-1 of Eimeria tenella.
XX
KW Poultry vaccine; Eimeria tenella vaccine; Eimeria necatrix vaccine.
XX OS Eimeria tenella.
XX
FH Key Location/Qualifiers
FT Peptide 24..181
FT Peptide /label=17,000 D peptide
FT Peptide 185..253
FT Peptide /label=8,000 D peptide
XX
PN A08665869-A.
XX
PD 04-JUN-1987.
XX
PF 01-DEC-1986; 86AU-0065869.
XX
PR 11-DEC-1985; 85US-0808013.
PR 03-DEC-1985; 85US-0805301.
PR 06-DEC-1985; 85US-0805824.
PR 11-DEC-1985; 85US-0807497.
XX
PA (SOLV ) SOLVAY & CIE.
XX
PI Newman KZ, Gore TC, Tedesco JL, Peterson GR, Brothers VM;
PI Paul LS, Chang RJ, Andrews WH; Files JG, Sias SR;
PI Kuhn I, McCaman MT, Files JG, Sias SR;
XX
DR WPI; 1987-199027/29.
DR N-PSDB; AAN70804.
XX
PT New polypeptide for protecting chickens against coccidiosis - are
PT obtd. by recombinant DNA procedures involving total genomic DNA
PT from eimeria maxima oocysts.
XX
PS Example; Fig 3; 255pp; English.
XX
CC Eimeria tenella polypeptides are used to immunise chickens against
CC infection against Eimeria strains, so that coccidiosis is prevent.
CC They are used in vaccines, eg in a dose to provide over 20
CC microgram/kg. The monoclonal antibody or its anti-idiotypic antibody
CC may also be administered to confer protection. The sequences in Fig
CC 3 (AAN70804, AAP70487) are very similar (identical apart from DNA
CC ambiguities?) to the sequences in Fig 5 (AAN70805, AAP70488).
XX
SQ Sequence 253 AA;
Query Match 32.4%; Score 44; DB 8; Length 253;
Best Local Similarity 40.0%; Pred. No. 1.1e+02;
Matches 10; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 3 SAITXQCKENLPISLCSXQGLRAEEN 27
Db 28 tavltdckeamnkirkgaaglpafed 52
:|:| ||| : | || | | :

RESULT 26
AAP70488
ID AAP70488 standard; protein; 253 AA.
XX
AC AAP70488;
XX
DT 02-APR-1991 (first entry)
XX
DE Sequence of the T4A antigen encoded by the BglII/EcoRI DNA fragment
DE of genomic clone 108-1 of Eimeria tenella.
XX
KW Poultry vaccine; Eimeria tenella vaccine; Eimeria necatrix vaccine.
XX OS Eimeria tenella.
XX
FH Key Location/Qualifiers
FT Peptide 24..181
FT Peptide /label=17,000 D peptide
FT Peptide 185..253
FT Peptide /label=8,000 D peptide
XX
PN A08665869-A.
XX
PD 04-JUN-1987.
XX
PF 01-DEC-1986; 86AU-0065869.
XX
PR 11-DEC-1985; 85US-0808013.
PR 03-DEC-1985; 85US-0805301.
PR 06-DEC-1985; 85US-0805824.
PR 11-DEC-1985; 85US-0807497.
XX
PA (SOLV ) SOLVAY & CIE.
XX
PI Newman KZ, Gore TC, Tedesco JL, Peterson GR, Brothers VM;
PI Paul LS, Chang RJ, Andrews WH; Files JG, Sias SR;
PI Kuhn I, McCaman MT, Files JG, Sias SR;
XX
DR WPI; 1987-199027/29.
DR N-PSDB; AAN70805.
XX
PT New polypeptide for protecting chickens against coccidiosis - are
PT obtd. by recombinant DNA procedures involving total genomic DNA
PT from eimeria maxima oocysts.
XX
PS Example; Fig 5; 255pp; English.
XX
CC Eimeria tenella polypeptides are used to immunise chickens against
CC infection against Eimeria strains, so that coccidiosis is prevent.
CC They are used in vaccines, eg in a dose to provide over 20
CC microgram/kg. The monoclonal antibody or its anti-idiotypic antibody
CC may also be administered to confer protection. The sequences in Fig
CC 3 (AAN70804, AAP70487) are very similar (identical apart from DNA
CC ambiguities?) to the sequences in Fig 5 (AAN70805, AAP70488).
XX
SQ Sequence 253 AA;
Query Match 32.4%; Score 44; DB 8; Length 253;
Best Local Similarity 40.0%; Pred. No. 1.1e+02;
Matches 10; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 3 SAITXQCKENLPISLCSXQGLRAEEN 27
Db 28 tavltdckeamnkirkgaaglpafed 52
:|:| ||| : | || | | :

```

RESULT 27

AAR32869
ID AAR32869 standard; Protein; 253 AA.

XX AC AAR32869;

XX DT 11-JUN-1993 (first entry)

XX DE Elmeria tenella pre-pro-TA4 antigen.

XX KW Coccidiosis; chicken; poultry; Apicomplexa; vaccine.

XX OS Elmeria tenella.

XX FH Key Location/Qualifiers

FT Peptide 1..23

FT Protein /label= signal

FT Protein /note= "17 kD protein"

FT Protein 185..253

FT Protein /note= "8 kD protein"

XX US5187080-A.

XX PD 16-FEB-1993.

XX PF 05-JUN-1984; 84US-0617483.

XX PR 05-JUN-1984; 84US-0617483.

XX PR 16-MAY-1985; 85US-0734085.

XX PR 06-DEC-1985; 85US-0805824.

XX PR 24-NOV-1987; 87US-0125012.

XX PR 21-JAN-1992; 92US-0824700.

XX PA (SOLV) SOLVAY & CIE.

XX PI Andrews WH, Brothers VM, Files JG, Gore TC, Kuhn I;

XX PI McCaman WT, Newman KZ, Paul LS, Sias SR, Tedesco JL;

XX DR WPI: 1993-075726/09.

XX DR N-PSDB; AAQ37481.

XX PT Nucleic acid encoding Eimeria tenella proteins - capable of inducing immune response in chickens, useful in vaccine

XX PS Example 6; Fig 5; 62pp; English.

XX CC This amino acid sequence was deduced from the nucleotide sequence of clone 108-1 isolated from an E. tenella sporulated oocyst genomic library. The clone hybridised strongly to one probe and weakly to two others; all the probes were based on amino acid sequences of the purified 17 kD component of the TA4 antigen. From the predicted amino acid sequence, it seems that the two peptides of the TA4 antigen (17 kD and 8 kD) are encoded by a contiguous nucleotide sequence, and at least one proteolytic step occurs to generate the 8 kD peptide by removal of a tripeptide. Both subunits are capable of inducing an immune response in chickens to confer protection against E. tenella.

XX SQ Sequence 253 AA;

Query Match

Best Local Similarity 32.4%; Score 44; DB 14; Length 253;

Matches 10; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 3 SAITXQCKENLPSLCGXQGLRAEN 27

Db 28 tavltdckeamnkirkaglpafed 52

RESULT 28

ABG23820

ID ABG23820 standard; Protein; 832 AA.

XX AC ABG23820;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #23811.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI: 2001-639362/73.

XX DR N-PSDB; AAS88007.

XX PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity

XX PS Claim 20; SEQ ID No 54179; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 832 AA;

Query Match

Best Local Similarity 32.4%; Score 44; DB 22; Length 832;

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 8 QCKENLPSL 16

Db 814 eckenlpsi 822

RESULT 29

ABG17156

ID ABG17156 standard; Protein; 932 AA.
 XX AC ABG17156;
 XX DT 18-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #17147.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX OS Homo sapiens.
 XX PN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX PA (HYSE-) HYSEQ INC.
 XX FI Drmanac RT, Liu C, Tang YT;
 XX DR WPI: 2001-639362/73.
 XX DR N-PSDB; AAS81343.
 XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX PS Claim 20; SEQ ID NO 47515; 103pp; English.
 XX CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 932 AA;
 Query Match 32.4%; Score 44; DB 22; Length 932;
 Best Local Similarity 77.8%; Pred. No. 5e+02;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Qy 8 OCKENLPSL 16
 :|||||:
 Db 799 eckenlpsi 807
 RESULT 30
 ABB58562
 ID ABB58562 standard; Protein; 1507 AA.

XX ABB58562;
 XX DT 26-MAR-2002 (first entry)
 XX DE Drosophila melanogaster polypeptide SEQ ID NO 2478.
 XX KW Drosophila; developmental biology; cell signalling; insecticide;
 XX KW pharmaceutical.
 XX OS Drosophila melanogaster.
 XX PN WO200171042-A2.
 XX PD 27-SEP-2001.
 XX PF 23-MAR-2001; 2001WO-US09231.
 XX PR 23-MAR-2000; 2000US-191637P.
 XX PR 11-JUL-2000; 2000US-0614150.
 XX PA (PEKE) PE CORP NY.
 XX FI Venter JC, Adams M, Li PWD, Myers EW;
 XX DR WPI: 2001-656860/75.
 XX DR N-PSDB; ABL02665.
 XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX PS Disclosure; SEQ ID NO 2478; 21pp + Sequence Listing; English.
 XX CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 1507 AA;
 Query Match 32.4%; Score 44; DB 22; Length 1507;
 Best Local Similarity 42.3%; Pred. No. 8.7e+02;
 Matches 11; Conservative 5; Mismatches 8; Indels 2; Gaps 1;
 Qy 4 AITXQCKE--NLPSCSXQGLRAEEN 27
 :||:|:|||||:|:
 Db 814 sitkspsedslpslcpnsitgdd 839
 RESULT 31
 AAG56627
 ID AAG56627 standard; Protein; 105 AA.
 XX AC AAG56627;
 XX DT 18-OCT-2000 (first entry)
 XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 72833.
 XX KW Protein identification; signal transduction pathway; metabolic pathway;
 XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
 XX KW termination sequence.
 XX OS Arabidopsis thaliana.

PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 31.6%; Score 43; DB 21; Length 105;

Best Local Similarity 50.0%; Pred. No. 56;

Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 8 QCKENLPSLCXQG 21

DB 21 qerrrlpricktlg 34

RESULT 32

AAG25073
ID AAG25073 standard; Protein; 153 AA.

XX AC AAG25073;

XX DT DT

XX DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 28989.

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

XX Arabidopsis thaliana.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX XX 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX XX

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.

DR N-PSDE; ABL10382.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 25629; 2lpp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 177 AA;

Query Match 31.6%; Score 43; DB 22; Length 177;
Best Local Similarity 47.1%; Pred. No. le+02;
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 11 ENLPISLCSXQGLRAEEN 27
| : | | | | | | |
Db 161 ekipclogagcgrrgtln 177

RESULT 34
AAG58013
ID AAG58013 standard; Protein; 208 AA.
AC AAG58013;
XX
XX 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 74834.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.

PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147460.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159233.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

```
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 31.6%; Score 43; DB 21; Length 210;
Best Local Similarity 50.08; Pred. No. 1.3e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Oy 11 ENLPSLCXQGLRAEE 26
   ||| :| || ||
Db 145 eslpricflsgltgee 160

RESULT 36
AAG55978
ID AAG55978 standard; Protein; 212 AA.
XX AC AAG55978;
XX AC AAG55978;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 71871.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 23-APR-1999; 99US-0130891.
XX PR 28-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 30-APR-1999; 99US-0132407.
XX PR 04-MAY-1999; 99US-0132484.
XX PR 05-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 06-MAY-1999; 99US-0132487.
XX PR 07-MAY-1999; 99US-0132863.
XX PR 11-MAY-1999; 99US-0134256.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 14-MAY-1999; 99US-0134370.
XX PR 18-MAY-1999; 99US-0134768.
XX PR 19-MAY-1999; 99US-0134941.
XX PR 20-MAY-1999; 99US-0135124.
XX PR 21-MAY-1999; 99US-0135353.
XX PR 24-MAY-1999; 99US-0135629.
XX PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
```

PR	04-AUG-1999;	99US-0147204.		PR	29-OCT-1999;	99US-0162142.	
PR	04-AUG-1999;	99US-0147302.			Query Match	31.6%;	Score 43; DB 21; Length 212;
PR	05-AUG-1999;	99US-0147192.			Best Local Similarity	50.0%;	Pred. No. 1.3e+02;
PR	05-AUG-1999;	99US-0147260.			Matches	8; Conservative	2; Mismatches 6; Indels 0; Gaps 0;
PR	06-AUG-1999;	99US-0147303.					
PR	06-AUG-1999;	99US-0147416.			QY	11 ENLPSLCSXOGLRAEE	26
PR	09-AUG-1999;	99US-0147493.				:	
PR	09-AUG-1999;	99US-0147935.			Db	147 eslpricflsgltgee	162
PR	10-AUG-1999;	99US-0148171.					
PR	11-AUG-1999;	99US-0148319.			RESULT	37	
PR	12-AUG-1999;	99US-0148341.			AAW60112		
PR	13-AUG-1999;	99US-0148365.			ID	AAW60112 standard; Protein; 326 AA.	
PR	13-AUG-1999;	99US-0148565.			XX		
PR	16-AUG-1999;	99US-0149368.			AC	AAW60112;	
PR	17-AUG-1999;	99US-0149175.			XX	25-AUG-1998 (first entry)	
PR	18-AUG-1999;	99US-0149426.			XX	Mycobacterium vaccae antigen 85A.	
PR	20-AUG-1999;	99US-0149722.			XX	Mycobacterium vaccae; antigen; therapy; prevention; cytokine production;	
PR	20-AUG-1999;	99US-0149723.			DE	M. avium; M. tuberculosis; immune response enhancer; cell proliferation;	
PR	20-AUG-1999;	99US-0149929.			XX	mycobacteria infection; vaccine; cancer.	
PR	23-AUG-1999;	99US-0149902.			XX	Mycobacterium vaccae.	
PR	23-AUG-1999;	99US-0149930.			OS	WO9808542-A2.	
PR	25-AUG-1999;	99US-0150566.			PN	05-MAR-1998.	
PR	26-AUG-1999;	99US-0150884.			PD		
PR	26-AUG-1999;	99US-0150884.			XX	28-AUG-1997; 97WO-NZ00105.	
PR	27-AUG-1999;	99US-0151065.			XX	12-JUN-1997; 97US-0873970.	
PR	27-AUG-1999;	99US-0151066.			PR	29-AUG-1996; 96US-0705347.	
PR	27-AUG-1999;	99US-0151080.			XX		
PR	30-AUG-1999;	99US-0151303.			PA	(GENE-) GENESIS RES & DEV CORP.	
PR	31-AUG-1999;	99US-0151438.			XX	Hiyama J, Prestidge RL, Scott LM, Skinner MA, Tan P;	
PR	01-SEP-1999;	99US-0151930.			PI	Visser E;	
PR	07-SEP-1999;	99US-0152363.			XX	WPI; 1998-216926/19.	
PR	10-SEP-1999;	99US-0153070.			DR	N-PSDB; AAV34569.	
PR	13-SEP-1999;	99US-0153758.			DR		
PR	15-SEP-1999;	99US-0154018.			XX	Mycobacterium vaccae polypeptides - used to develop products for use	
PR	16-SEP-1999;	99US-0154039.			PT	in detection, therapy and prevention of mycobacteria infections or	
PR	20-SEP-1999;	99US-0154779.			PT	as immune response enhancers	
PR	22-SEP-1999;	99US-0155139.			XX	Example 4; Pages 81-82; 153pp; English.	
PR	23-SEP-1999;	99US-0155486.			XX	This represents a Mycobacterium vaccae antigen 85A. The invention	
PR	24-SEP-1999;	99US-0155659.			CC	provides M. vaccae polypeptides that comprise an immunogenic portion of	
PR	28-SEP-1999;	99US-0156458.			CC	a soluble M. vaccae antigen, or a variant, where the antigen induces an	
PR	29-SEP-1999;	99US-0156596.			CC	immune response in patients previously exposed to a mycobacterium. Such	
PR	04-OCT-1999;	99US-0157117.			CC	M. vaccae polypeptides can be used in methods for enhancing non-specific	
PR	05-OCT-1999;	99US-0157753.			CC	immune response. The methods and products can be used for the detection,	
PR	06-OCT-1999;	99US-0157865.			CC	treatment and prevention of infectious diseases caused by mycobacteria	
PR	07-OCT-1999;	99US-0158029.			CC	such as M. vaccae, M. avium or M. tuberculosis. The products also have	
PR	08-OCT-1999;	99US-0158232.			CC	the ability to induce cell proliferation and cytokine production (e.g.	
PR	12-OCT-1999;	99US-0158369.			CC	interferon-gamma and interleukin-12 production) in T cells, NK cells,	
PR	13-OCT-1999;	99US-0159293.			CC	B cells, or macrophages. They can be used for enhancing immune responses	
PR	13-OCT-1999;	99US-0159294.			CC	for use in vaccines or immunotherapy of infectious diseases and cancers.	
PR	13-OCT-1999;	99US-0159295.			XX	SQ	Sequence 326 AA;
PR	14-OCT-1999;	99US-0159329.					
PR	14-OCT-1999;	99US-0159330.				Query Match	31.6%;
PR	14-OCT-1999;	99US-0159331.				Best Local Similarity	52.9%;
PR	14-OCT-1999;	99US-0159637.				Matches	9; Conservative
PR	14-OCT-1999;	99US-0159638.					3; Mismatches 5; Indels 0; Gaps 0;
PR	18-OCT-1999;	99US-0159584.				QY	11 ENLPSLCSXOGLRAEEN
PR	21-OCT-1999;	99US-0160741.					:
PR	21-OCT-1999;	99US-0160767.				Db	71 enspalyllldgraged
PR	21-OCT-1999;	99US-0160768.					87
PR	21-OCT-1999;	99US-0160770.					
PR	21-OCT-1999;	99US-0160814.					
PR	21-OCT-1999;	99US-0160815.					
PR	22-OCT-1999;	99US-0160980.					
PR	22-OCT-1999;	99US-0160981.					
PR	22-OCT-1999;	99US-0160989.					
PR	25-OCT-1999;	99US-0161404.					
PR	25-OCT-1999;	99US-0161405.					
PR	25-OCT-1999;	99US-0161406.					
PR	26-OCT-1999;	99US-0161359.					
PR	26-OCT-1999;	99US-0161360.					
PR	26-OCT-1999;	99US-0161361.					
PR	28-OCT-1999;	99US-0161920.					
PR	28-OCT-1999;	99US-0161992.					
PR	28-OCT-1999;	99US-0161993.					

RESULT 38	RESULT 39
AA114858	AA08670
ID AA114858 standard; protein; 326 AA.	ID AA08670 standard; Protein; 344 AA.
XX	XX
AC AA114858;	AC AA08670;
XX	XX
DT 25-OCT-1999 (first entry)	DT 17-OCT-2000 (first entry)
XX	XX
DE M. vaccae antigen 85A.	DE Arabidopsis thaliana protein fragment SEQ ID NO: 6300.
XX	XX
KW Mycobacterium vaccae protein; antigen; T cell activation; cytokine;	KW Protein identification; signal transduction pathway; metabolic pathway;
KW dendritic cell maturation; infectious disease; immune disorder; cancer;	KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW respiratory system; mycobacterial infection; allergy; tuberculosis;	KW termination sequence.
KW leprosy; sarcoidosis; lung cancer; asthma; skin disorder; psoriasis;	OS Arabidopsis thaliana.
KW dermatitis; eczema; alopecia areata; skin cancer; basal carcinoma;	XX
KW squamous cell carcinoma; melanoma.	PN EP1033405-A2.
XX	XX
OS Mycobacterium vaccae.	PD 06-SEP-2000.
XX	XX
PN WO9932634-A2.	XX
XX	XX
PD 01-JUL-1999.	PF 25-FEB-2000; 2000EP-0301439.
XX	XX
PF 23-DEC-1998; 98WO-NZ00189.	PR 25-FEB-1999; 99US-0121825.
XX	PR 05-MAR-1999; 99US-0123180.
XX	PR 09-MAR-1999; 99US-0123548.
PR 04-DEC-1998; 98US-0205426.	PR 23-MAR-1999; 99US-0125788.
PR 23-DEC-1997; 97US-0996624.	PR 25-MAR-1999; 99US-0126284.
PR 23-DEC-1997; 97US-0997080.	PR 29-MAR-1999; 99US-0126785.
PR 23-DEC-1997; 97US-0997362.	PR 01-APR-1999; 99US-0127462.
PR 11-JUN-1998; 98US-0095855.	PR 08-APR-1999; 99US-0128234.
PR 17-SEP-1998; 98US-0156181.	PR 16-APR-1999; 99US-0129845.
XX	PR 19-APR-1999; 99US-0130077.
PA (GENE-) GENESIS RES & DEV CORP LTD.	PR 21-APR-1999; 99US-0130449.
XX	PR 23-APR-1999; 99US-0130510.
XX	PR 23-APR-1999; 99US-0130891.
PI Prestidge RL, Skinner MA, Tan P, Visser ES, Watson J;	PR 28-APR-1999; 99US-0131449.
XX	PR 30-APR-1999; 99US-0132048.
DR WPI; 1999-430163/36.	PR 30-APR-1999; 99US-0132407.
DR N-PSDB; AAZ11303.	PR 04-MAY-1999; 99US-0132484.
XX	PR 05-MAY-1999; 99US-0132485.
PT Enhancing immune response to an antigen	PR 06-MAY-1999; 99US-0132486.
XX	PR 06-MAY-1999; 99US-0132487.
PS Example 12; Page 162-163; 243pp; English.	PR 07-MAY-1999; 99US-0132863.
XX	PR 11-MAY-1999; 99US-0134256.
CC The invention provides heat-killed Mycobacterium vaccae, or recombinant	PR 14-MAY-1999; 99US-0134218.
CC M. vaccae proteins. The M. vaccae proteins may be employed to activate	PR 14-MAY-1999; 99US-0134219.
CC T cells and natural killer cells, to stimulate the production of	PR 14-MAY-1999; 99US-0134221.
CC cytokines, to enhance the expression of co-stimulatory molecules on	PR 18-MAY-1999; 99US-0134370.
CC dendritic cells and monocytes, and to enhance dendritic cell maturation	PR 19-MAY-1999; 99US-0134768.
CC and function. The proteins can be expressed by standard recombinant	PR 19-MAY-1999; 99US-0134941.
CC methodology. Pharmaceutical compositions comprising the proteins or	PR 20-MAY-1999; 99US-0135124.
CC nucleic acid sequences encoding the proteins can be used for the	PR 21-MAY-1999; 99US-0135353.
CC treatment, prevention, and detection of disorders including infectious	PR 24-MAY-1999; 99US-0135629.
CC diseases, immune disorders and cancer. In particular, the compounds and	PR 25-MAY-1999; 99US-0136021.
CC methods are used for treatment of diseases of the respiratory system,	PR 27-MAY-1999; 99US-0136392.
CC such as mycobacterial infections, asthma, allergies, tuberculosis,	PR 28-MAY-1999; 99US-0136782.
CC leprosy, sarcoidosis and lung cancers, and disorders of the skin such as	PR 01-JUN-1999; 99US-0137222.
CC psoriasis, atopic dermatitis, eczema, allergic contact dermatitis,	PR 03-JUN-1999; 99US-0137528.
CC alopecia areata, and skin cancers such as basal carcinoma, squamous cell	PR 04-JUN-1999; 99US-0137502.
CC carcinoma and melanoma.	PR 07-JUN-1999; 99US-0137724.
XX	PR 08-JUN-1999; 99US-0138094.
SQ Sequence 326 AA;	PR 10-JUN-1999; 99US-0138540.
	PR 10-JUN-1999; 99US-0138847.
	PR 14-JUN-1999; 99US-0139119.
Query Match 31.68; Score 43; DB 20; Length 326;	PR 16-JUN-1999; 99US-0139452.
Best Local Similarity 52.99; Pred. NO. 2.1e+02;	PR 16-JUN-1999; 99US-0139453.
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;	PR 17-JUN-1999; 99US-0139492.
QY 11 ENLPSCSXOGLRAEEN 27	PR 18-JUN-1999; 99US-0139454.
: :	PR 18-JUN-1999; 99US-0139455.
Db 71 enspalyldglraqed 87	PR 18-JUN-1999; 99US-0139456.
	PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 31.6%; Score 43; DB 21; Length 344;

Best Local Similarity 41.2%; Pred. No. 2.2e+02;

Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 12 NLPISLCSXQGLRAENI 28

Db 229 nfpiclhsgmsqeel 245

RESULT 40

AAU57026

ID AAU57026 standard; Protein; 361 AA.

XX

AC AAU57026;

XX

DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #17922.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 XX 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 XX (CORI-) CORIXA CORP.
 XX
 XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI; 2001-616774/71.
 DR N-PSDB; AAS59579.
 XX
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 PS Example 1; SEQ ID No 18221; 1069pp; English.
 XX
 CC Sequences AAU39105-AAUG8017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 361 AA;
 Query Match 31.6%; Score 43; DB 22; Length 361;
 Best Local Similarity 38.1%; Pred. No. 2.4e+02;
 Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;
 Qy 2 GSAITXOCKENPLSLCSXQGL 22
 ||| : : | | | :
 Db 264 gsavvakreglaflctalgi 284

Search completed: August 22, 2002, 07:50:24
 Job time: 187 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 07:49:02 ; Search time 41.85 Seconds
(without alignments)
119.877 Million cell updates/sec

Title: US-09-856-199-1

Perfect score: 164

Sequence: 1 KRCINQLCKLPTDSELAPRSXCCHRH 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

SPTREMBL_19.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	125	76.2	386	6 Q9N2H6	Q9n2h6 bos taurus
2	122	74.4	380	11 Q9Z2T7	Q9z2t7 mus musculus
3	53	32.3	168	11 Q9D9I2	Q9d9i2 mus musculus
4	53	32.3	194	11 Q9QY08	Q9qy08 rattus norv
5	53	32.3	417	11 Q35520	Q35520 rattus norv
6	53	32.3	559	11 Q63135	Q63135 rattus norv
7	53	32.3	1292	10 Q9LDN0	Q9ldn0 arabidopsis
8	53	32.3	1662	10 Q23274	Q23274 arabidopsis
9	52	31.7	363	5 Q9VI35	Q9vi35 drosophila
10	52	31.7	1698	5 Q94438	Q94438 chironomus
11	51.5	31.4	1698	11 Q9C264	Q9c264 mus musculus
12	51.5	31.4	518	5 Q76403	Q76403 caenorhabdi
13	51	31.1	549	5 Q9U3F1	Q9u3f1 caenorhabdi
14	51	31.1	553	5 Q9GTC3	Q9gtc3 caenorhabdi
15	51	31.1	556	5 Q9XXU7	Q9xxu7 caenorhabdi
16	51	31.1	685	5 Q9VXL1	Q9vxl1 drosophila

17	51	31.1	694	5 Q9GZ15	Q9gz15 drosophila
18	51	31.1	970	4 Q9SL07	Q9sl07 homo sapien
19	50.5	30.8	772	10 Q9SJ14	Q9sj14 arabidopsis
20	50	30.5	617	13 Q90499	Q90499 echis pyram
21	49	29.9	71	10 Q9LRE0	Q9lre0 brassica ca
22	49	29.9	845	13 Q9DGR1	Q9dgr1 xenopus lae
23	49	29.9	1300	11 Q9WTL4	Q9wtl4 mus musculus
24	49	29.9	4706	10 Q9FN44	Q9fn44 arabidopsis
25	48	29.3	129	5 Q9UAH3	Q9uah3 pinctada fu
26	48	29.3	225	5 Q9NAD8	Q9nad8 caenorhabdi
27	48	29.3	313	10 Q9LX00	Q9lxx0 arabidopsis
28	48	29.3	318	10 Q949L4	Q949l4 cicier ariet
29	48	29.3	320	2 Q56299	Q56299 thriobacillu
30	48	29.3	558	10 Q9M084	Q9m084 arabidopsis
31	48	29.3	1042	5 P90974	P90974 caenorhabdi
32	48	29.3	1333	11 Q9Z0U5	Q9z0u5 rattus norv
33	48	29.3	1516	4 Q12764	Q12764 homo sapien
34	47.5	29.0	201	13 Q9PTC3	Q9ptc3 notechis at
35	47.5	29.0	209	5 Q9W3F3	Q9w3f3 drosophila
36	47.5	29.0	1779	5 Q9VW49	Q9vw49 drosophila
37	47	28.7	173	11 Q9D4K2	Q9d4k2 mus musculus
38	47	28.7	195	2 Q9ETN5	Q9etn5 streptococc
39	47	28.7	228	5 Q18836	Q18836 caenorhabdi
40	47	28.7	318	10 Q949L3	Q949l3 cicier ariet
41	47	28.7	332	5 Q9GQ80	Q9gqb0 brachiosteo
42	47	28.7	483	13 Q9IAB0	Q9iab0 agkistrodon
43	47	28.7	505	4 Q9NVU2	Q9nvu2 homo sapien
44	47	28.7	591	10 Q9FVQ9	Q9fvg9 arabidopsis
45	47	28.7	789	12 Q06120	Q06120 potato viru
46	47	28.7	1722	5 Q19350	Q19350 caenorhabdi
47	47	28.7	3059	12 Q9QBT7	Q9qbt7 potato viru
48	47	28.7	3059	12 Q9QBT6	Q9qbt6 potato viru
49	47	28.7	3059	12 Q9QBT5	Q9qbt5 potato viru
50	47	28.7	3059	12 Q9DIC2	Q9dic2 potato viru
51	47	28.7	3059	12 Q85197	Q85197 potato viru
52	46.5	28.4	103	12 Q912B6	Q912b6 human eryth
53	46.5	28.4	380	1 Q50515	Q50515 methanother
54	46.5	28.4	383	17 Q27210	Q27210 methanother
55	46.5	28.4	2342	12 Q65980	Q65980 cherry capi
56	46.5	28.4	2809	4 Q96JP8	Q96jp8 homo sapien
57	46	28.0	309	11 Q9WV09	Q9wv09 mus musculu
58	46	28.0	398	8 Q34939	Q34939 leishmania
59	46	28.0	509	4 Q96CJ0	Q96cj0 homo sapien
60	46	28.0	660	5 Q9VCR3	Q9vcr3 drosophila
61	46	28.0	736	3 Q74399	Q74399 schizosacch
62	46	28.0	750	11 Q922L1	Q922l1 mus musculu
63	46	28.0	817	5 Q9W5L0	Q9w5l0 drosophila
64	46	28.0	1030	4 Q9BY59	Q9bys9 homo sapien
65	46	28.0	1245	3 Q9Y7V5	Q9y7v5 trichoderma
66	45.5	27.7	201	13 Q9PU35	Q9pu35 notechis sc
67	45.5	27.7	201	13 Q9PU34	Q9pu34 notechis sc
68	45.5	27.7	202	13 Q9PU33	Q9pu33 notechis sc
69	45.5	27.7	1017	5 Q9VHL7	Q9vhl7 drosophila
70	45	27.4	140	12 Q9WR16	Q9wr16 african cas
71	45	27.4	165	10 Q94JV9	Q94jv9 arabidopsis
72	45	27.4	172	5 Q9NIV2	Q9niv2 drosophila
73	45	27.4	175	10 Q23368	Q23368 arabidopsis
74	45	27.4	271	2 Q93G14	Q93g14 marinobacte
75	45	27.4	274	5 Q9VH65	Q9vh65 drosophila
76	45	27.4	293	4 Q99944	Q99944 homo sapien
77	45	27.4	434	10 Q9SJ16	Q9sj16 arabidopsis
78	45	27.4	457	10 Q94LK1	Q94lk1 oryza sativ
79	45	27.4	569	10 Q93YP6	Q93yp6 arabidopsis
80	45	27.4	705	4 Q9H624	Q9h624 homo sapien
81	45	27.4	716	2 Q9EUP1	Q9eup1 corynebacte
82	45	27.4	814	4 Q9NY78	Q9ny78 homo sapien
83	45	27.4	902	10 Q9M364	Q9m364 arabidopsis
84	45	27.4	930	4 Q9NY77	Q9ny77 homo sapien
85	45	27.4	1125	4 Q9NY75	Q9ny75 homo sapien
86	45	27.4	1156	4 Q95938	Q95938 homo sapien
87	45	27.4	1176	4 Q9NY76	Q9ny76 homo sapien
88	45	27.4	1193	10 Q9MA43	Q9ma43 arabidopsis
89	45	27.4	1827	4 Q9H482	Q9h482 homo sapien

90	45	27.4	2117	4	Q9H4D8
91	45	27.4	2150	5	Q4131
92	45	27.4	2167	4	Q9H481
93	45	27.4	2169	4	Q9NY09
94	45	27.4	3058	12	Q9QBT4
95	45	27.4	3229	5	Q26912
96	44.5	27.1	121	5	Q9H4Z4
97	44.5	27.1	200	11	Q9D169
98	44.5	27.1	202	13	Q9PTC2
99	44.5	27.1	340	4	Q9HG60
100	44.5	27.1	501	12	Q91GL2
					Q9H4d8 homo sapien
					Q4131 caenorhabdi
					Q9H481 homo sapien
					Q9ny09 homo sapien
					Q9qbt4 potato viru
					Q26912 trypanosoma
					Q9h4z4 caenorhabdi
					Q9d169 mus musculus
					Q9ptc2 notechis at
					Q9h6m0 homo sapien
					Q91gl2 epiphyas po

ALIGNMENTS

```

RESULT      1
Q9N2H6      PRELIMINARY;      PRF;      386 AA.
ID          Q9N2H6
AC          Q9N2H6;
DT          01-OCT-2000 (TtEMBLrel. 15, Created)
DT          01-OCT-2000 (TtEMBLrel. 15, Last sequence update)
DT          01-OCT-2000 (TtEMBLrel. 15, Last sequence update)
DE          SELENOPROTEIN P.
GN          SLP.
OS          Bos taurus (Bovine).
OC          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC          Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC          Bovidae; Bovinae; Bos.
OC          NCBI_TaxID=9913;
RN          [1]
RP          SEQUENCE FROM N.A.
RA          Hara S., Imura N., Shoji Y.;
RT          Bovine endothelial selenoprotein P.";
RL          Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR          EMBL; AB032826; BAA84781.1.
SQ          SEQUENCE      386 AA; 43061 MW;      EC6F3A76874E528F CRC64;

```

```

RESULT      2
Q922T7      PRELIMINARY;      PRRT;      380 AA.
ID          AC
Q922T7;
DT          01-MAY-1999 (TREMBlrel. 10, Created)
DT          01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT          01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DT          01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE          PLASMA SELENOPROTEIN P.
GN          SELP.
OS          Mus musculus (Mouse).
OC          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX          NCBI_taxid=10090;
[1]
RN          SEQUENCE FROM N.A.
RP
RC          STRAIN=129;
RX          MEDLINE=97434516; PubMed=9288402;
RT          Steinert P., Ahrens M., Gross G., Flohe L.;
RT          "cDNA and deduced polypeptide sequence of a mouse selenoprotein P.";
RL          BioFactors 6:311-319(1997).
RL          EMBL; AF021345; AAC01684.1;
SQ          SEQUENCE      380 AA; 42375 MW; C5055C18D32BDD14 CRC64;

```

Best Local Similarity	77.8%;	Pred. No.	3.1e-11;		
Matches	21;	Conservative	1;	Mismatches	5;
				Indels	0;
				Gaps	0;
QY	3	CINOLCKLPTDSELAPRSCXCHRH	29		
			:		
DB	281	CINOLCKLKESEAPSSCCCHRH	307		
RESULT	3				
Q9D9I2		PRELIMINARY;	PRT;	168	AA.
ID	Q9D9I2				
AC	Q9D9I2;				
DT	01-JUN-2001	(TEMBLrel. 17, Created)			
DT	01-JUN-2001	(TEMBLrel. 17, Last sequence update)			
DT	01-JUN-2001	(TEMBLrel. 17, Last annotation update)			
DE	4931420D14RIK	PROTEIN.			
GN	4931420D14RIK.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RC	SEQUENCE FROM N.A.				
RP	SPRAIN=C57BL/6J; TISSUE=TESTIS;				
RX	MEDLINE=21085660; PubMed=11217851;				
RA	Kawal J., Shingagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,				
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,				
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,				
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,				
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,				
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,				
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,				
RA	Schirrali L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,				
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,				
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,				
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,				
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,				
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,				
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,				
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,				
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,				
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,				
RA	Hayashizaki Y.				
RT	"Functional annotation of a full-length mouse cDNA collection."				
RL	Nature 409:685-690(2001).				
DR	EMBL; AK006892; BAB24782.1; -.				
MG	MG; MGI:1913992; 4931420D14RIK.				
SQ	SEQUENCE 168 AA; 18931 MW; 7A2BD279612A5E94 CRC64;				

Query Match 74.48; Score 122; DB 11; Length 380;

Matches	8;	Conservative	3;	Mismatches	5;	Indels	0;	Gaps	0;
QY	10	KLPTDSELA PRSXCCH 25							
Db	552	RLPEDSDVAMEDACCH 567							
RESULT	8								
O23274									
AC	O23274	PRELIMINARY;	PRT;	1662	AA.				
DT	01-JAN-1998	(TREMBlrel. 05, Created)							
DT	01-JAN-1998	(TREMBlrel. 05, Last sequence update)							
DT	01-JUN-2001	(TREMBlrel. 17, Last annotation update)							
DE	KINESIN HOMOLOG (KINESIN LIKE PROTEIN).								
GN	AT4G14150.								
OS	Arabidopsis thaliana (Mouse-ear cross).								
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;								
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;								
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi.								
OX	NCBI_TaxID=3702;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RX	MEDLINE=9812113; PubMed=9461215;								
RA	Bevan M., Bancroft I., Bent E., Love K., Goodman H., Dean C.,								
RA	Bergkamp R., Dirke W., van Staveren M., Stiekema W., Drost L.,								
RA	Ridley P., Hudson S.A., Patel K., Murphy G., Piffanelli P., Wedler H.,								
RA	Wedler E., Wambutt R., Weitzenger T., Pohl T.M., Terry N.,								
RA	Gielen J., Villarroel R., De Clerck R., van Montagu M., Lecharny A.,								
RA	Auborg S., Gy I., Kreis M., Lao N., Kavanagh T., Hempel S., Kotter P.,								
RA	Autvin K.D., Rieger M., Schaeffer M., Funk B., Mueller-Auer S.,								
RA	Silvey M., James R., Montfort A., Pons A., Puigdomenech P., Douka A.,								
RA	Voukelatos E., Milioni D., Hatzopoulos P., Piravandi E., Obermaier B.,								
RA	Hilbert H., Duesterhoft A., Moores T., Jones J.D.G., Eneva T.,								
RA	Palme K., Benes V., Rechman S., Ansoorge W., Cooke R., Berger C.,								
RA	Delsenly M., Voet M., Volckaert G., Mewes H.W., Klosterman S.,								
RA	Schueler C., Chalwatzis N.								
RT	"Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of								
RT	Arabidopsis thaliana."								
RL	Nature 391:485-488(1998).								
RN	[2]								
RP	SEQUENCE FROM N.A.								
RA	EU Arabidopsis sequencing project;								
RL	Submitted (MAR-2000) to the EMBL/GenBank/DDJB databases.								
DR	EMBL; Z97335; CAB10194.1; -.								
DR	EMBL; AL161538; CAB78457.1; -.								
DR	HSSP; p17119; 3KAR.								
DR	InterPro; IPR001752; kinesin.								
DR	Pfam; PF00225; kinesin; 2.								
DR	PRINTS; PR00380; KINESINHEAVY.								
DR	SMART; SM00129; KISC; 1.								
DR	PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.								
DR	PROSITE; PS0067; KINESIN_MOTOR_DOMAIN2; 1.								
KW	ATP-binding; Coiled coil; Microtubules; Motor protein.								
SQ	SEQUENCE 1662 AA; 186876 MW; D006426D84FAC4AD CRC64;								
Query Match	32.3%;	Score 53;	DB 10;	Length 1662;					
Best Local Similarity	50.0%;	Pred. No. 10;							
Matches	8;	Conservative	3;	Mismatches	5;	Indels	0;	Gaps	0;
QY	10	KLPTDSELA PRSXCCH 25							
Db	543	RLPEDSDVAMEDACCH 558							
RESULT	9								
O9VI35									
AC	O9VI35	PRELIMINARY;	PRT;	363	AA.				
DT	01-MAY-2000	(TREMBlrel. 13, Created)							
DT	01-MAY-2000	(TREMBlrel. 13, Last sequence update)							
DT	01-JUN-2001	(TREMBlrel. 17, Last annotation update)							

DE	CG2337	PROTEIN.							
GN	CG2337.								
OS	Drosophila melanogaster (Fruit fly).								
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;								
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;								
OC	Ephydroidea; Drosophilidae; Drosophila.								
OX	NCBI_TaxID=7227;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=BERKELEY;								
RX	MEDLINE=20196006; PubMed=10731132;								
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,								
RA	Amantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,								
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,								
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,								
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,								
RA	Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.D.G.,								
RA	April J.F., Adayani A., An H.-J., Andrews-Frannkoch C., Baldwin D.,								
RA	Balieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,								
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,								
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,								
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,								
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,								
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,								
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,								
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,								
RA	Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,								
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,								
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,								
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,								
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,								
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,								
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,								
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,								
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,								
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,								
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,								
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,								
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,								
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,								
RA	Splier E., Spradling A.C., Stapleton M., Strong R., Sun E.,								
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,								
RA	Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,								
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,								
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,								
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,								
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;								
RT	"The genome sequence of Drosophila melanogaster."								
RL	Science 287:2185-2195(2000).								
DR	EMBL; AE003674; AAF54109.1; -.								
DR	HSSP; P04284; 1CFE.								
DR	FlyBase; FBgn0037457; CG2337.								
DR	InterPro; IPR001283; SCP.								
DR	Pfam; PF00188; SCP; 1.								
DR	PRINTS; PR00837; V5TFPLIKE.								
DR	SMART; SM00198; SCP; 1.								
SQ	SEQUENCE 363 AA; 40688 MW; C67A5E99953B8182 CRC64;								
Query Match	31.7%;	Score 52;	DB 5;	Length 363;					
Best Local Similarity	22.2%;	Pred. No. 3.4;							
Matches	6;	Conservative	10;	Mismatches	11;	Indels	0;	Gaps	0;
QY	3	CINQLLKLPTDSELA PRSXCCHRL 29							
Db	136	CLCEVICQYPNAKSMKPQASCOHNKQI 162							
RESULT	10								
O94438									
ID	O94438	PRELIMINARY;	PRT;	1698	AA.				
AC	O94438								
DT	01-FEB-1997	(TREMBlrel. 02, Created)							

DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE 185 KDA SILK PROTEIN.
GN SPI85.
OS Chironomus pallidivittatus (Midge).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomidae; Chironominae; Chironomus.
OX NCBI_TaxID=7151;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-SALIVARY GLAND;
RA Case S.T., Cox C., Bell W.C., Hoffman R.T., Martin J., Hamilton R.;
RT "Extraordinary conservation of cysteines among homologous Chironomus
silk proteins spi85 and spi220."
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U54640; AAA99803.1; -
DR InterPro: IPR004153; CXCXC.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000853; Mclthion_nemat.
DR Pfam: PF03128; CXCXC; 71.
DR PRINTS: PR00876; MTNEMAT0E.
DR PROSITE: PS00022; EGF-1; UNKNOWN.1.
SQ SEQUENCE 1698 AA; 186164 MW; 82AF0838D9E2360E CRC64;

Query Match 31.7%; Score 52; DB 5; Length 1698;
Best Local Similarity 41.7%; Pred. No. 15;
Matches 10; Conservative 1; Mismatches 11; Indels 2; Gaps 1;

OY 3 CINQLLCKLPDSELA PRSKXCHC 26
I:| | | | | | | | | |
Db 777 CVNPPKCNSP--QVLKDESCCHC 798

RESULT 11

Q9CZ64 AC Q9CZ64 PRELIMINARY; PRT; 169 AA.
ID 09CZ64
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzei J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Suzuki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Sasaki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL: AK012967; BAB28573.1; -.

DR MGD; MGI:1923803; 2810402K13RIK.
SQ SEQUENCE 169 AA; 18777 MW; 402A9BF9292579FD CRC64;

Query Match 31.4%; Score 51.5; DB 11; Length 169;
Best Local Similarity 37.0%; Pred. No. 1.9;
Matches 10; Conservative 4; Mismatches 12; Indels 1; Gaps 1;

OY 1 KRCINOLLCK-LPTDSELA PRSKXCHC 26
I:| | | | | | | | | |
Db 38 KRAVMSRAEGLPRSLQFAQLYCCCH 64

RESULT 12

O76403 AC O76403 PRELIMINARY; PRT; 518 AA.
ID 076403
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE F36F12.1 PROTEIN.
GN F36F12.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fulton L.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans."
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Smith A., Harmon G.;
RT "The sequence of C. elegans cosmid F36F12."
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF067940; AAC19200.1; -
DR InterPro: IPR002875; DUF23.
DR Pfam: PF01697; DUF23; 1.
DR ProDom: PD004153; DUF23; 1.
SQ SEQUENCE 518 AA; 59646 MW; 6161E8CD0109A851 CRC64;

Query Match 31.4%; Score 51.5; DB 5; Length 518;
Best Local Similarity 33.3%; Pred. No. 5.7;
Matches 9; Conservative 5; Mismatches 6; Indels 7; Gaps 1;

OY 2 RCINOLLCKLPDSELA PRSKXCHCRH 28
I:| | | | | | | | | |
Db 467 QCVNTKACKLPNDKIP-----CRH 486

RESULT 13

Q9U3F1 AC Q9U3F1 PRELIMINARY; PRT; 549 AA.
ID Q9U3F1

QVXL1;
01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DT CG9163 PROTEIN.
GN MMD OR CG9163
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=BERKELEY;
RC MEDLINE=20196006; PubMed=10731132;
RX ADAMAN M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Adams M.D., Celniker S.E., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA Ananides P.G., Scher S.E., Richards S., Ashburner M., Henderson S.N.,
RA George R.A., Lewis S.E., Yandell M.D., Zhang Q., Chen L.X.,
RA Sutton G.G., Wortman J.R., Venter A., White O., Holt R.A.,
RA Brandon K.C., Rogers J.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berland P.P., Bhattacharya D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hartis N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwac K.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003500; AAF48548.1; -.
DR HSSP: P18619; 1FVL.
DR MEROPS: M12.UNB; -.
DR FlyBase: FBgn0041109; mmd.
DR InterPro: IPRO01762; Disintegrin.
DR InterPro: IPRO00561; EGF-like.
DR InterPro: IPRO02870; Pep_M12B_propep.
DR InterPro: IPRO01590; Reprolysin.
DR Pfam: PF00200; disintegrin; 1.
DR Pfam: PF01562; Pep_M12B_propep; 1.
DR Pfam: PF01421; Reprolysin; 1.
DR PRINTS: PR00289; DISINTEGRIN.
DR ProDom: PD000664; Disintegrin; 1.
DR SMART: SM00050; DISIN; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS50215; ADAM_MEPRO; 1.
DR PROSITE: PS50214; DISINTEGRIN_2; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
SQ SEQUENCE 685 AA; 75625 MW; 02AC9C506578CF3 CRC64;

Query Match 31.1%; Score 51; DB 5; Length 685;
Best Local Similarity 30.8%; Pred. NO. 9;
Matches 8; Conservative 8; Mismatches 10; Indels 0; Gaps

QY 2 RCINOLLKLPDSELAAPRSXCCHCR 27
DB 402 QCDCGITCKLKSEAQCAGACCDCQR 387

RESULT 17
Q9GZ15 PRELIMINARY; PRT; 694 AA.

ID AC Q9GZ15;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
MIND-MELD (FRAGMENT).
GN MMD OR CG9163.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
[1]
RN SEQUENCE FROM N.A.
RP Harper C.H., Chase B.A.;
RT "The Drosophila mind-meld gene encodes a neuronally expressed ADAM protein.";
RL Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AF252287; AAF98331.1; -.
DR HSSP: P18619; 1FVL.
DR MEROPS: M12.UNB; -.
DR FlyBase: FBgn0041109; mmd.
DR InterPro: IPRO01762; Disintegrin.
DR InterPro: IPRO00561; EGF-like.
DR InterPro: IPRO02870; Pep_M12B_propep.
DR InterPro: IPRO01590; Reprolysin.
DR Pfam: PF00200; disintegrin; 1.
DR Pfam: PF01562; Pep_M12B_propep; 1.
DR Pfam: PF01421; Reprolysin; 1.
DR PRINTS: PR00289; DISINTEGRIN.
DR ProDom: PD000664; Disintegrin; 1.
DR SMART: SM00050; DISIN; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS50215; ADAM_MEPRO; 1.
DR PROSITE: PS50214; DISINTEGRIN_2; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 694 AA; 76317 MW; 37736F7B064A72E9 CRC64;

Query Match 31.1%; Score 51; DB 5; Length 694;
Best Local Similarity 30.8%; Pred. NO. 9;
Matches 8; Conservative 8; Mismatches 10; Indels 0; Gaps

QY 2 RCINOLLKLPDSELAAPRSXCCHCR 27
DB 362 QCDCGITCKLKSEAQCAGACCDCQR 387

RESULT 18
Q9PI07 PRELIMINARY; PRT; 970 AA.

ID AC Q9PI07;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE GEM-INTERACTING PROTEIN.
GN GMIP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

SEQUENCE FROM N.A.
 RA Aresta S., Berger F., Berger R., de Gunzburg J.;
 RT "GMP, a Gem interacting protein.",
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF132541; AAF61330.1; -
 DR HSSP; Q07960; IRGP.
 DR InterPro; IPR002086; Aldehyde dehydr.
 DR InterPro; IPR002219; DAG_PE-bind.
 DR InterPro; IPR000198; RhOGAP.
 DR SMART; SM00109; Cl; 1.
 DR SMART; SM00324; RhOGAP; 1.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; UNKNOWN_1.
 DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
 SQ SEQUENCE 970 AA; 106732 MW; 3A1E35AF348F6949 CRC64;

Query Match 31.1%; Score 51; DB 4; Length 970;
 Best Local Similarity 57.9%; Pred. No. 12;
 Matches 11; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

QY 6 QLLCKLPTDSELAPRSXC 24
 :||| || ||| ||| ||| |||
 Db 942 RLLSKL--DSEAVPRATCC 958

RESULT 19

ID Q9SJ14 PRELIMINARY; PRT; 772 AA.
 AC Q9SJ14;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE AT2G21850 PROTEIN.
 GN AT2G21850.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]

SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.";
 RL Nature 402:761-768(1999).
 RN [2]

SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC007019; AAD20422.1; -
 DR InterPro; IPR002219; DAG_PE-bind.
 DR InterPro; IPR004146; DC1.
 DR InterPro; IPR001965; PHD.
 DR InterPro; IPR001841; ZnF_ring.
 DR Pfam; PF03107; DC1; 1.
 DR SMART; SM00109; Cl; 5.
 DR SMART; SM00249; PHD; 4.
 DR SMART; SM00184; RING; 2.
 DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 2.
 SQ SEQUENCE 772 AA; 88510 MW; EE9B74E1CF63D772 CRC64;

Query Match 30.8%; Score 50.5; DB 10; Length 772;
 Best Local Similarity 40.0%; Pred. No. 12;
 Matches 14; Conservative 3; Mismatches 7; Indels 11; Gaps 2;

QY 1 KRCI-----NQLCKLPTDSELAPRSXCCHCRH 28
 ||| : |||: || | : | | | | |
 Db 131 KRCPGPDARGSCLLCELP----LSPSSVCYGCYV 161

RESULT 20

ID Q90499 PRELIMINARY; PRT; 617 AA.
 AC Q90499;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE METALLOPROTEASE.
 GN ECHI.
 OS Echis pyramidum (carpet viper).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 OC Viperidae; Viperinae; Echis.
 OX NCBI_TaxID=8700;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=VENOM GLAND;
 RX MEDLINE=95010025; PubMed=7925363;
 RA Paine M.J.I., Moura-da-Silva A.M., Theakston R.D.G., Crampton J.M.;
 RT "Cloning of metalloprotease genes in the carpet viper (Echis pyramidum leakeyi): Further members of the metalloprotease/disintegrin gene family.";
 RL Eur. J. Biochem. 224:483-488(1994).
 CC -1- SIMILARITY: BELONGS TO THE DISINTEGRIN FAMILY.
 DR EMBL; X78970; CAA55565.1; -
 DR HSSP; P18619; IFVL.
 DR MEROPS; M12.0PB; -
 DR InterPro; IPR001762; Disintegrin.
 DR InterPro; IPR002870; Pep_M12B_propep.
 DR InterPro; IPR001590; Repolysin.
 DR InterPro; IPR000130; Zn_MTpeptdse.
 DR Pfam; PF00200; disintegrin; 1.
 DR Pfam; PF01562; Pep_M12B_propep; 1.
 DR Pfam; PF01421; Repolysin; 1.
 DR PRINTS; PR00289; DISINTEGRIN.
 DR ProDom; PD000664; Disintegrin; 1.
 DR SMART; SM00050; DISIN; 1.
 DR PROSITE; PS50215; ADAM_MEPRO; 1.
 DR PROSITE; PS00427; DISINTEGRIN_1; 1.
 DR PROSITE; PS50214; DISINTEGRIN_2; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
 KW Protease.
 FT CHAIN 193 617 METALLOPROTEASE.
 SQ SEQUENCE 617 AA; 69310 MW; 83DC3DA5F4F3AF8A CRC64;

Query Match 30.5%; Score 50; DB 13; Length 617;
 Best Local Similarity 40.0%; Pred. No. 12;
 Matches 10; Conservative 1; Mismatches 14; Indels 0; Gaps 0;

QY 3 CINQLCKLPTDSELAPRSXCCHCR 27
 | | | | | | | | | | | | |
 Db 430 CCNATTKLTGSCQADGECNCQR 454

RESULT 21

ID Q9LRE0 PRELIMINARY; PRT; 71 AA.
 AC Q9LRE0;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)

DE S LOCUS PROTEIN 11-32 (FRAGMENT).
GN SP11-32.
OS Brassica campestris (Field mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3711;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S32(50-11T);
RX MEDLINE=20273842; PubMed=10812061;
RA Watanabe M., Ito A., Takada Y., Ninomiya C., Kakizaki T., Takahata Y.,
RA Hatakeyama K., Hinata K., Suzuki G., Takasaki T., Satta Y., Shiba H.,
RA Takayama S., Isogai A.;
RT "Highly divergent sequences of the pollen self-incompatibility (S)
RT gene in class-I S haplotypes of Brassica campestris (syn. rapa) L.";
RL FEBS Lett. 473:139-144(2000).
DR EMBL; AB039756; BAA96394.1; -.
FT NON_TER
SQ SEQUENCE 71 AA; 7928 MW; 1E8E5E88D8720F97 CRC64;

Query Match 29.9%; Score 49; DB 10; Length 71;
Best Local Similarity 35.7%; Pred. No. 2.1;
Matches 10; Conservative 4; Mismatches 10; Indels 4; Gaps 1;

QY 1 KRCINOLLCKLPTDSELAPRSXCCHRH 28
:||: : : : |||
DB 35 QRCVTEFSKKIDK----PRLSCICRH 58

RESULT 22
ID Q9DGR1 PRELIMINARY; PRT; 845 AA.
AC Q9DGR1
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HOMOLOG OF HUMAN MT-SPI.
GN XMT-SPI.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20363741; PubMed=10903452;
RA Yamada K., Takabatake T., Takeshima K.;
RT "Isolation and characterization of three novel serine protease genes
RT from Xenopus laevis";
RL Gene 252:209-216(2000).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL; AB038498; BAB08218.1; -.
DR HSSP; P00763; 1DFO.
DR MEROPS; S01.050; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR002172; LDL_recept_A.
DR InterPro; IPR001254; Trypsin.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00057; ldl_recept_a; 4.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00192; LDLa; 3.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS50068; LDLa_2; 4.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN.1.
KW Glycoprotein; Hydrolase; Serine protease.
SQ SEQUENCE 845 AA; 93597 MW; 7FD7E62851A758B7 CRC64;

Query Match 29.9%; Score 49; DB 13; Length 845;
Best Local Similarity 45.5%; Pred. No. 22;
Matches 10; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 2 RCINOLLCKLPTDSELAPRSXC 23
:||| : : : |||
DB 759 RIINQTECNKLLDGLTPRMLC 780

RESULT 23
ID Q9WTL4 PRELIMINARY; PRT; 1300 AA.
AC Q9WTL4
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE INSULIN RECEPTOR-RELATED RECEPTOR.
GN INSR OR IRR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99272238; PubMed=10342810;
RA Hirayama I., Tamemoto H., Yokota H., Kubo S.-K., Wang J., Kuwano H.,
RA Nagamachi Y., Takeuchi T., Izumi T.;
RT "Insulin receptor-related receptor is expressed in pancreatic b-cells
RT and stimulates tyrosine phosphorylation of insulin receptor substrate-
RT 1 and-2";
RL Diabetes 48:1237-1244(1999).
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP + PROTEIN
CC TYROSINE PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AB007135; BAA77835.1; -.
DR HSSP; P06213; IIRK.
DR MGD; MGI:1346037; Insrr.
DR InterPro; IPR000494; EGFR_L.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR002011; Receptor_tyr_kin_II.
DR InterPro; IPR000531; TonB_boxC.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR PRINTS; PR00109; TYRKINASE.
DR SMART; SM00060; FN3; 3.
DR SMART; SM00261; FU; 1.
DR SMART; SM00219; TyrkC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN.1.
KW ATP-binding; Glycoprotein; Phosphorylation; Receptor; Transferase;
KW Transmembrane; Tyrosine-protein kinase.
SQ SEQUENCE 1300 AA; 144745 MW; 29B9FC3C521A2067 CRC64;

Query Match 29.9%; Score 49; DB 11; Length 1300;
Best Local Similarity 32.3%; Pred. No. 34;
Matches 10; Conservative 2; Mismatches 7; Indels 12; Gaps 1;

QY 10 KLPTDS-----ELAPRSXCCHRH 28
:||| : : : |||

Db 662 RLPTSHDTRFDREDPALEAPEQGCCCCQH 692

```
RESULT 24
Q9FN44
ID Q9FN44 PRELIMINARY; PRT; 4706 AA.
AC Q9FN44;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE GENOMIC DNA, CHROMOSOME 5, PI CLONE: MJ24.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RX MEDLINE=98069011; PubMed=9405937;
RA Kotani H., Nakamura Y., Sato S., Kaneko T., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. II.
RT Sequence features of the regions of 1,044,062 bp covered by thirteen
RT physically assigned pl clones."
RL DNA Res. 4:291-300(1997).
CC -|- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL: AB006708; BAB09828.1; -
DR InterPro: IPR001841; Znf_ring.
DR Pfam: PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
KW Zinc-finger.
SQ SEQUENCE 4706 AA; 527237 MW; BB78EF17227E5C35 CRC64;
```

Query Match 29.9%; Score 49; DB 10; Length 4706;
Best Local Similarity 47.4%; Pred. No. 1.2e+02;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 8 LCKLPTDSELPAPRSCXCHC 26
||: ||| ||| |||
Db 3771 LCIISGDSFLKPYDECCFC 3789

```
RESULT 25
Q9UAH3
ID Q9UAH3 PRELIMINARY; PRT; 129 AA.
AC Q9UAH3;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE N14#2.PRO.
GS N14#2.SEQ.
OS Pinctada fucata.
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pterioidea;
OC Pterioidea; Pterilioidea; Pinctada.
OX NCBI_TaxID=50426;
RN [1]
RP SEQUENCE FROM N.A.
RA Hayashi N., Samata T.;
RT "14 kDa matrix protein family in nacreous layer of Japanese pearl
RT oyster, Pinctada fucata."
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB023249; BAA83734.1; -
SQ SEQUENCE 129 AA; 15375 MW; DIC569125B687604 CRC64;
```

Query Match 29.3%; Score 48; DB 5; Length 129;
Best Local Similarity 34.4%; Pred. No. 5.4;
Matches 11; Conservative 3; Mismatches 14; Indels 4; Gaps 1;

QY 1 KRCINOLLCKLPTDSELPAPR----SXCHCRH 28
|:| | : | | | | | | | | | |

Db 27 KCGRYSYCWIPYDIERRDRNGKKCCFCRN 58

```
RESULT 26
Q9NAD8
ID Q9NAD8 PRELIMINARY; PRT; 225 AA.
AC Q9NAD8;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE Y51H4A.15 PROTEIN.
GS Y51H4A.15.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Sulston J.E.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL: AL132952; CAB61147.1; -
SQ SEQUENCE 225 AA; 25965 MW; A4B7E976973EE778 CRC64;
```

Query Match 29.3%; Score 48; DB 5; Length 225;
Best Local Similarity 71.4%; Pred. No. 9.1;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 INQLCKLPTDSEL 17
| | | | | | | | | |
Db 155 IQLLSKLPSDEEL 168

```
RESULT 27
Q9LXQ0
ID Q9LXQ0 PRELIMINARY; PRT; 313 AA.
AC Q9LXQ0;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2001 (TRENBLrel. 18, Last annotation update)
DE HYPOTHETICAL 36.4 KDA PROTEIN.
GS F26G5_80.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA D'Angelo M., Vezzi A., Modesto D., Pigazzi M., Valle G., Mewes H.W.,
RA Rudd S., Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL353814; CAB88421.1; -
DR InterPro: IPR001810; F-box.
DR Pfam: PF00646; F-box; 1.
DR SMART; SM00256; FBOX; 1.
DR PROSITE; PS0181; FBOX; 1.
KW Hypothetical protein.
SQ SEQUENCE 313 AA; 36439 MW; EE553779AC20EB14 CRC64;
```

Query Match 29.3%; Score 48; DB 10; Length 313;

Best Local Similarity 45.0%; Pred. NO. 13;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 6 OLLCKLPTDSELAPRSXCH 25
Db 13 EILCRPLGSLVRLRSVKH 32

RESULT 28

Q949L4 ID Q949L4 PRELIMINARY; PRT; 318 AA.

AC Q949L4;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PUTATIVE POLYPEPTIDE (FRAGMENT).
GN POL.
OS Cicer arietinum (Chickpea) (Garbanzo).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae; Cicer.
OX NCBI_TaxID=3827;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON-TY3-GP5Y LIKE RETROTRANSPONSON CAREP;
RA Staginuss C.;
RL Thesis (2001), Department of Plant Molecular Biology,
RL University of Frankfurt / M., Frankfurt a.M., Germany.
DR EMBL: AJ411799; CAC51363.1;
KW Polyprotein.
FT NON_TER 1 1
FT 318 318
SQ SEQUENCE 318 AA; 35294 MW; 5E106DC2B1CD3ADE CRC64;

Query Match 29.3%; Score 48; DB 10; Length 318;
Best Local Similarity 34.3%; Pred. NO. 13;
Matches 12; Conservative 3; Mismatches 10; Indels 10; Gaps 2;

QY 3 CINQL---LCKLP-----TDELAPRSXCHCR 27
Db 174 CYNRLKLSYALPFDLVSTPKTLTVNSACLHCR 208

RESULT 29

Q56299 ID Q56299 PRELIMINARY; PRT; 320 AA.

AC Q56299;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PUTATIVE REPLICATION PROTEIN.
OS Thiobacillus intermedius.
OG Plasmid pTik12.
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Thiomonas.
OX NCBI_TaxID=926;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE=95273471; PubMed=7753904;
RA English R.S., Lorbach S.C., Hufman K.M., Shively J.M.;
RT "Isolation and characterization of the replicon of a Thiobacillus
intermedius plasmid."
RL plasmid 33:1-6(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RA Shively J.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: U36865; AAC12677.1;
DR InterPro: IPR004322; Replicase.
DR Pfam: PF03090; Replicase; 1.
KW Plasmid.

SQ SEQUENCE 320 AA; 36202 MW; F858CB0AE80916F CRC64;

Query Match 29.3%; Score 48; DB 2; Length 320;
Best Local Similarity 53.3%; Pred. NO. 13;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 13 TDELAPRSXCHCR 27
Db 295 TESRCAPRAACNRR 309

RESULT 30

Q9M084 ID Q9M084 PRELIMINARY; PRT; 558 AA.

AC Q9M084;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ASPARTATE--TRNA LIGASE-LIKE PROTEIN.
GN ATG31180.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lennard N., Quail M., Harris B., Rajandream M.A., Barrell B.G.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + L-AMINO ACID + TRNA(AMINO ACID) = AMP +
CC -!- PYROPHOSPHATE + L-AMINOACYL-TRNA(AMINO ACID).
CC -!- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
DR EMBL: AL161578; CAB79836.1;
DR HSP; P04802; 1ASZ.
DR InterPro: IPR002106; AA_trna_ligase_II.
DR InterPro: IPR004364; trna-synt_2.
DR InterPro: IPR002312; trna-synt_2.
DR InterPro: IPR002313; trna-synt_lys_2.
DR InterPro: IPR004365; trna-anti.
DR Pfam: PF00152; trna-synt_2; 1.
DR Pfam: PF01336; trna-anti; 1.
DR PRINTS; PR01042; TRNASYNTHASP.
DR PROSITE; PS00179; AA_TRNA_LIGASE_II_1; 1.
DR PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
KW ATP-binding; Aminoacyl-TRNA synthetase; Ligase; Protein biosynthesis.

Query Match 29.3%; Score 48; DB 10; Length 558;
Best Local Similarity 57.9%; Pred. NO. 22;
Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 3 CINQLCKLPDSELAPRS 21
Db 196 CINLSLAKPLSVSDAARS 214

RESULT 31

P90974 ID P90974 PRELIMINARY; PRT; 1042 AA.

AC P90974;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ADM-1 PREPROTEIN PRECURSOR.
GN ADM-1 OR Y37D8A.13.
OS Caenorhabditis elegans.

DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DE 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE KIAA0191 PROTEIN (FRAGMENT).
 GN KIAA0191.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nagase T., Seki N., Ishikawa K., Tanaka A., Nomura N.;
 FT "Prediction of the coding sequences of unidentified human genes. V.
 RT The coding sequences of 40 new genes (KIAA 0161 - KIAA 0200) deduced
 RT by analysis of cDNA clones from human cell line KG-1.";
 RL DNA Res. 0:0-0(0).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96281124; PubMed=8724849;
 RA Nagase T., Seki N., Ishikawa K., Tanaka A., Nomura N.;
 RT "Prediction of the coding sequences of unidentified human genes. V.
 RT The coding sequences of 40 new genes (KIAA0161-KIAA0200) deduced by
 RT analysis of cDNA clones from human cell line KG-1.";
 RL EMBL; D83776; BAAL2105.1; -;
 DR InterPro; IPR002934; NTP_transf.
 DR InterPro; IPR001201; PAP_25A_core.
 DR InterPro; IPR002058; PAP_assoc.
 DR InterPro; IPR000822; znf-C2H2.
 DR InterPro; IPR001878; znf-CCHC.
 DR Pfam; PF01909; NTP_transf_2; 1.
 DR Pfam; PF00098; znf-CCHC; 3.
 DR SMART; SM00343; znf-C2HC; 3.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 KW Zinc-finger.
 FT NON_TER
 SQ SEQUENCE 1516 AA; 171110 MW; 3FB7736E4FB8AE26 CRC64;
 1
 Query Match 29.3%; Score 48; DB 4; Length 1516;
 Best Local Similarity 66.7%; Pred. No. 56;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 13 TUSELAPRSXCC 24
 II IIII II
 DB 1156 TDGELAPNDRCC 1167
 PRELIMINARY; PRT; 201 AA.
 RESULT 34
 Q9PTC3
 AC Q9PTC3
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DE 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE PHOSPHOLIPASE A2 INHIBITOR ALPHA SUBUNIT ISOFORM NAI-2A
 DE PRECURSOR.
 OS Notochis ater.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroidae;
 OC Elapidae; Acanthophinae; Notochis.
 OX NCBI_TaxID=111176;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sekuloski S., Dunn R.D., Broady K.W.;
 FT "Identification of alpha and beta subunit isoforms of a phospholipase
 RT A2 inhibitor isolated from four species of Elapidae.";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF211162; AAF23779.1; -;
 DR InterPro; IPR000345; CytC_heme_bind.
 DR InterPro; IPR001526; Lf6_UPAR.
 DR InterPro; IPR004126; PLA2_inh.
 DR Pfam; PF02988; PLA2_inh; 1.

DR SMART; SM00134; LU; 1.
 DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
 KW Signal.
 FT SIGNAL
 GN KIAA0191
 FT CHAIN 1 19
 20 201
 PHOSPHOLIPASE A2 INHIBITOR ALPHA SUBUNIT
 FT ISOFORM NAI-2A.
 SQ SEQUENCE 201 AA; 22315 MW; FBE89DB7FE665744 CRC64;
 29.0%; Score 47.5; DB 13; Length 201;
 Best Local Similarity 35.5%; Pred. No. 9.8; Indels 7; Gaps 1;
 Matches 11; Conservative 3; Mismatches 10;
 QY 1 KRINQLLCKLPT-----DSELAPRSXCC 24
 ::I: IIII II I I I
 DB 66 RCESSSLKLERFDINIGHDSYLRGRTHCC 96
 RESULT 35
 Q9W3F3
 ID Q9W3F3 PRELIMINARY; PRT; 209 AA.
 AC Q9W3F3
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DE 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
 DE CG15347 PROTEIN.
 GN CG15347.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrell J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
 RA Balleg R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kimons I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";

```

RL Science 287:2185-2195(2000).
DR EMBL; AE003444; AAF46374.1; -.
DR FlyBase; FBgn0030040; CG15347.
SQ SEQUENCE 209 AA; 22798 MW; F4D3DF1292A668DB CRC64;

Query Match 29.08; Score 47.5; DB 5; Length 209;
Best Local Similarity 38.78; Pred. No. 10;
Matches 12; Conservative 3; Mismatches 9; Indels 7; Gaps 1;

Oy 3 CINQLLCKLPDSELAAPRSX-----CCHC 26
   I I I I I I I I I I I
Db 72 CNNELLCQCTYSEGCNRNTEPFLSLMCLQC 102

RESULT 36
Q9VM49 PRELIMINARY; PRT; 1779 AA.
AC Q9VM49;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CG8334 PROTEIN (FRAGMENT).
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbavani A., An H.-J., Andrews-pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman D.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burdick K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matti B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
RT Science 287:2185-2195(2000).
CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
DR EMBL; AE003515; AAF49100.2; -.

```

```

DR MEROPS; C19.UPW; -.
DR FlyBase; FBgn0036913; CG8334.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001230; Prenyltn.
DR InterPro; IPR001125; Recoverin.
DR InterPro; IPR001394; UCH-2.
DR Pfam; PF00036; efhand; 2.
DR Pfam; PF00442; UCH-1; 1.
DR Pfam; PF00443; UCH-2; 1.
DR PRINTS; PR00450; RECOVERIN.
DR SMART; SM00054; EFH; 2.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.
DR PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
DR PROSITE; PS00972; UCH_2_1; 1.
DR PROSITE; PS00973; UCH_2_2; 1.
DR PROSITE; PS00973; UCH_2_3; 1.
DR PROSITE; PS00235; UCH_2_3; 1.
KW Calcium-binding.
FT NON_TER 1
SQ SEQUENCE 1779 AA; 199146 MW; 8062479CA02FA652 CRC64;

Query Match 29.08; Score 47.5; DB 5; Length 1779;
Best Local Similarity 50.08; Pred. No. 79;
Matches 14; Conservative 2; Mismatches 7; Indels 5; Gaps 2;

Oy 5 NQLLCKLP-----TDSELAP-RSXGCHCR 27
   I I I I I I I I I I I I I
Db 1083 NQVHELPLDEAGKDSLELAQRNCCYSR 1110

RESULT 37
Q9D4K2 PRELIMINARY; PRT; 173 AA.
AC Q9D4K2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 4931420D14RIK PROTEIN.
GN 4931420D14RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyszaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK016467; BAB30253.1; -.
DR MGD; MGI:1913992; 4931420D14RIK.
SQ SEQUENCE 173 AA; 19581 MW; 147B6F155AC29FDF CRC64;

Query Match 28.78; Score 47; DB 11; Length 173;

```

Best Local Similarity 37.0%; Pred. No. 10;
Matches 10; Conservative 2; Mismatches 15; Indels 0; Gaps 0;

QY 1 KRCINQLCKLPTDSELAPRSXCCHCR 27

Db 47 KRSSHLLCGLSRSCCCRCRCYCR 73

RESULT 38

Q9ETN5

ID Q9ETN5 PRELIMINARY; PRT; 195 AA.

AC Q9ETN5

DT 01-MAR-2001 (TRENBLrel. 16, Created)

DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)

DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

DE ABC TRANSPORTER (PUTATIVE ABC TRANSPORTER BLPAORF1).

GN IA.

OS Streptococcus pneumoniae.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=1313;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=KNR7/87;

RA Reichmann P., Hakenbeck R.;

RT "A Peptide Inducible Signal Transduction System in Streptococcus

pneumoniae: Evidence for Bacteriocin Production.";

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA de Saizieu A., Gardes C., Flint N., Wagner C., Kamber M.,

RA Mitchell T.J., Keck W., Amrein K.E., Lange R.;

RT "Microarray based identification of a novel Streptococcus pneumoniae

regulon controlled by an autoinduced peptide.";

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ278419; CAC18586.1; -;

DR EMBL; AJ276410; CAC03519.1; -;

DR MEROPS; C39.001; -;

SQ SEQUENCE 195 AA; 22271 MW; A45B91733EC50314 CRC64;

Query Match

Best Local Similarity 28.7%; Score 47; DB 2; Length 195;

Matches 10; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 5 NOLLCKLPTDSELAPRSXCCHC 26

Db 152 NGLLSKLPSDFQTKISHCLHC 173

RESULT 39

Q18836

ID Q18836 PRELIMINARY; PRT; 228 AA.

AC Q18836; Q19856;

DT 01-NOV-1996 (TRENBLrel. 01, Created)

DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)

DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

DE C54D10.10 PROTEIN.

GN C54D10.10.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RA Kershaw J.;

RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=94150718; PubMed=7906398;

RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,

RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fulton L.,

RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,

RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kinston J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,
RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";

RL Nature 368:32-38(1994).

RN [3]

RP SEQUENCE FROM N.A.

RA Dobson R.;

RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; Z755339; CAA99845.1; -;

DR EMBL; Z755331; CAA99845.1; JOINED.

DR EMBL; Z755331; CAA99808.1; -;

DR EMBL; Z755339; CAA99808.1; JOINED.

DR HSP; P12111; 2KNT.

DR InterPro; IPR002899; EB.

DR InterPro; IPR002223; Kunitz_BPTI.

DR Pfam; PF00014; Kunitz_BPTI; 2.

DR PRINTS; PR00759; BASICPTASE.

DR SMART; SM00131; KU; 2.

DR SMART; SM00289; WRI; 1.

DR PROSITE; PS02079; BPTI_KUNITZ_2; 2.

KW Serine protease inhibitor.

SQ SEQUENCE 228 AA; 25884 MW; B7C8C1FB71CC06DB CRC64;

Query Match

Best Local Similarity 28.7%; Score 47; DB 5; Length 228;

Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 9 CKLPTDSELAPRSXCC 24

Db 133 CKLPTDAKIVTRAQKC 148

RESULT 40

Q949L3

ID Q949L3 PRELIMINARY; PRT; 318 AA.

AC Q949L3;

DT 01-DEC-2001 (TRENBLrel. 19, Created)

DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

DE PUTATIVE POLYPROTEIN (FRAGMENT).

GN POL.

OS Cicer arietinum (Chickpea) (Garbanzo).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae; Cicer.

OX NCBI_TaxID=3827;

RN [1]

RP SEQUENCE FROM N.A.

RC TRANSPOSON=TY3-GYPSY LIKE RETROTRANSPOSON CAREP;

RA Staginuss C.;

RL Thesis (2001), Department of Plant Molecular Biology,

University of Frankfurt / M., Frankfurt a.M., Germany.

DR EMBL; AJ411800; CAC44140.1; -;

KW Polyprotein.

FT NON_TER 1 1

FT NON_TER 318 318

SQ SEQUENCE 318 AA; 35253 MW; B70D99F5E4B48650 CRC64;

Query Match

Best Local Similarity 28.7%; Score 47; DB 10; Length 318;

Matches 12; Conservative 3; Mismatches 10; Indels 10; Gaps 2;

QY 3 CINQL---LCKLPTD-----SELAPRSXCCHCR 27

Db 174 CYNRLKLSVSLPFDLVVSTPAKTLTVNSACLHCR 208

Search completed: August 22, 2002, 07:52:14
Job time: 192 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 11:17:39 ; Search time 21.52 Seconds
(without alignments)
31.781 Million cell updates/sec

Title: US-09-856-199-2

Perfect score: 28

Sequence: 1 TGSATFXCKENLPISCSXQGLRAENI 28

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1200 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	21.4	443	4	US-09-088-425-2
2	6	21.4	508	4	US-08-980-523-9
3	6	21.4	553	1	US-07-683-957B-3
4	6	21.4	575	1	US-07-683-957B-2
5	6	21.4	685	3	US-08-947-965-74
6	6	21.4	2293	4	US-09-368-590-2
7	5	17.9	10	1	US-07-683-957B-4
8	5	17.9	10	1	US-08-447-411-16
9	5	17.9	10	1	US-08-447-411-48
10	5	17.9	10	2	US-08-662-227-5
11	5	17.9	10	4	US-09-017-947-5
12	5	17.9	13	5	PCT-US95-04121-3
13	5	17.9	21	1	US-08-447-411-20
14	5	17.9	21	1	US-08-447-411-60
15	5	17.9	21	2	US-08-662-227-17
16	5	17.9	21	4	US-09-017-947-17
17	5	17.9	22	1	US-08-447-411-73
18	5	17.9	22	2	US-08-662-227-31
19	5	17.9	22	4	US-09-017-947-31
20	5	17.9	30	1	US-08-145-708A-4
21	5	17.9	30	2	US-08-331-454-4
22	5	17.9	47	4	US-09-187-789-43
23	5	17.9	88	4	US-08-936-165A-369
24	5	17.9	96	1	US-08-417-460-10
25	5	17.9	102	4	US-08-477-347-9
26	5	17.9	106	2	US-08-850-910A-30
27	5	17.9	109	4	US-09-460-384-34
28	5	17.9	124	3	US-08-630-172-4
29	5	17.9	124	4	US-09-375-419-4
30	5	17.9	126	1	US-08-170-360-3
31	5	17.9	126	2	US-08-888-497-38
32	5	17.9	126	4	US-09-362-230-38
33	5	17.9	126	5	PCT-US94-07926-38
34	5	17.9	137	3	US-08-804-180C-2
35	5	17.9	148	2	US-08-888-497-36
36	5	17.9	148	4	US-09-362-230-36
37	5	17.9	148	5	PCT-US94-07926-36
38	5	17.9	158	4	US-08-679-493A-24
39	5	17.9	161	2	US-08-360-608B-33
40	5	17.9	174	3	US-08-804-180C-4
41	5	17.9	187	1	US-08-067-684-14
42	5	17.9	187	1	US-08-008-898-14
43	5	17.9	187	2	US-08-459-818-14
44	5	17.9	187	2	US-08-889-666-14
45	5	17.9	187	2	US-08-465-078-14
46	5	17.9	187	2	US-08-725-776-14
47	5	17.9	187	2	US-08-488-062-14
48	5	17.9	187	3	PCT-US95-06726-36
49	5	17.9	187	3	US-08-228-208A-14
50	5	17.9	195	2	US-08-882-704A-2
51	5	17.9	198	5	PCT-US91-01861-5
52	5	17.9	199	1	US-08-132-405-5
53	5	17.9	199	1	US-08-395-939A-5
54	5	17.9	199	4	US-09-430-503-42
55	5	17.9	199	4	US-09-430-503-44
56	5	17.9	199	4	US-09-430-503-46
57	5	17.9	199	4	US-09-430-503-48
58	5	17.9	207	4	US-08-446-114A-21
59	5	17.9	209	4	US-09-430-503-18
60	5	17.9	209	4	US-09-430-503-20
61	5	17.9	209	4	US-09-430-503-22
62	5	17.9	209	4	US-09-430-503-24
63	5	17.9	219	2	US-08-855-140-3
64	5	17.9	219	2	US-08-807-044-3
65	5	17.9	219	5	PCT-US91-04986-2
66	5	17.9	220	4	US-09-247-373B-50
67	5	17.9	221	4	US-09-247-373B-54
68	5	17.9	223	3	US-08-228-208A-17
69	5	17.9	234	1	US-08-505-058-1
70	5	17.9	234	1	US-08-505-058-2
71	5	17.9	234	2	US-08-459-818-21
72	5	17.9	234	2	US-08-459-818-22
73	5	17.9	234	2	US-08-889-666-21
74	5	17.9	234	2	US-08-889-666-22
75	5	17.9	234	2	US-08-465-078-21
76	5	17.9	234	2	US-08-465-078-22
77	5	17.9	234	2	US-08-465-078-22
78	5	17.9	234	2	US-08-725-776-21
79	5	17.9	234	2	US-08-725-776-22
80	5	17.9	234	2	US-08-488-062-21
81	5	17.9	234	2	US-08-488-062-22
82	5	17.9	243	2	US-08-829-110-3
83	5	17.9	264	1	US-08-485-618-93
84	5	17.9	264	1	US-08-362-652-93
85	5	17.9	264	2	US-08-605-672-93
86	5	17.9	264	2	US-08-482-293A-93
87	5	17.9	264	2	US-08-943-363-93
88	5	17.9	264	4	US-09-193-043-93
89	5	17.9	269	4	US-09-430-503-2
90	5	17.9	269	4	US-09-430-503-4
91	5	17.9	269	4	US-09-430-503-6
92	5	17.9	269	4	US-09-430-503-8
93	5	17.9	284	2	US-08-766-439-32
94	5	17.9	284	2	US-08-766-439-41
95	5	17.9	284	2	US-08-766-439-42
96	5	17.9	284	2	US-08-766-439-43
97	5	17.9	284	2	US-08-766-439-44
98	5	17.9	284	2	US-08-766-439-45
99	5	17.9	286	1	US-08-358-117-2
100	5	17.9	288	2	US-08-160-524A-6

101	5	17.9	322	2	US-08-892-880-2	Sequence 2, Appli	174	5	17.9	816	1	US-08-190-802A-54	Sequence 54, Appl
102	5	17.9	326	2	US-08-306-511A-10	Sequence 10, Appl	175	5	17.9	816	4	US-08-477-346-54	Sequence 54, Appl
103	5	17.9	326	2	US-08-893-274-10	Sequence 10, Appl	176	5	17.9	816	4	US-08-473-089-54	Sequence 54, Appl
104	5	17.9	326	3	US-08-581-918A-10	Sequence 10, Appl	177	5	17.9	822	4	US-09-564-805-222	Sequence 222, App
105	5	17.9	326	4	US-08-346-147B-10	Sequence 10, Appl	178	5	17.9	826	4	US-09-564-805-2	Sequence 2, Appli
106	5	17.9	326	4	US-08-822-936-10	Sequence 10, Appl	179	5	17.9	826	4	US-09-564-805-224	Sequence 224, App
107	5	17.9	326	5	PCT-US95-04636-10	Sequence 10, Appl	180	5	17.9	826	4	US-09-564-805-226	Sequence 226, App
108	5	17.9	344	5	US-08-681-192-2	Sequence 2, Appli	181	5	17.9	830	3	US-08-804-439A-20	Sequence 20, Appl
109	5	17.9	351	3	US-08-591-685-7	Sequence 7, Appli	182	5	17.9	830	3	US-08-720-229-20	Sequence 20, Appl
110	5	17.9	357	3	US-08-630-172-20	Sequence 20, Appl	183	5	17.9	853	4	US-09-354-129-10	Sequence 10, Appl
111	5	17.9	357	4	US-08-375-419-20	Sequence 20, Appl	184	5	17.9	871	2	US-08-775-009-34	Sequence 34, Appl
112	5	17.9	390	2	US-08-873-093-3	Sequence 3, Appli	185	5	17.9	885	3	US-08-310-912A-2	Sequence 2, Appli
113	5	17.9	396	1	US-08-539-798-3	Sequence 3, Appli	186	5	17.9	885	3	US-08-841-089-2	Sequence 2, Appli
114	5	17.9	396	1	US-08-329-560-3	Sequence 3, Appli	187	5	17.9	885	4	US-09-301-085-2	Sequence 2, Appli
115	5	17.9	396	1	US-08-471-791-16	Sequence 16, Appl	188	5	17.9	885	5	PCT-US95-04570-2	Sequence 2, Appli
116	5	17.9	396	5	PCT-US91-01746-16	Sequence 16, Appl	189	5	17.9	885	5	PCT-US95-04589-2	Sequence 2, Appli
117	5	17.9	401	2	US-08-805-118-1	Sequence 1, Appli	190	5	17.9	890	4	US-09-342-648-10	Sequence 10, Appl
118	5	17.9	401	4	US-09-391-958-1	Sequence 1, Appli	191	5	17.9	907	3	US-08-930-996A-7	Sequence 7, Appli
119	5	17.9	409	2	US-08-743-130A-2	Sequence 2, Appli	192	5	17.9	909	2	US-08-310-912A-142	Sequence 142, App
120	5	17.9	409	2	US-08-743-130A-39	Sequence 39, Appl	193	5	17.9	909	4	US-09-301-085-142	Sequence 142, App
121	5	17.9	414	5	PCT-US92-06840-2	Sequence 2, Appli	194	5	17.9	909	5	PCT-US95-04589-142	Sequence 142, App
122	5	17.9	432	3	US-09-075-087-2	Sequence 2, Appli	195	5	17.9	921	1	US-08-188-582-2	Sequence 2, Appli
123	5	17.9	432	4	US-09-472-971-1	Sequence 1, Appli	196	5	17.9	921	1	US-08-646-715-2	Sequence 2, Appli
124	5	17.9	450	2	US-08-861-464-2	Sequence 2, Appli	197	5	17.9	956	1	US-08-185-232A-2	Sequence 2, Appli
125	5	17.9	450	2	US-08-396-001-2	Sequence 2, Appli	198	5	17.9	956	1	US-08-416-523-2	Sequence 2, Appli
126	5	17.9	450	4	US-09-323-433A-2	Sequence 2, Appli	199	5	17.9	956	3	US-08-789-478-2	Sequence 2, Appli
127	5	17.9	453	3	US-08-993-359-22	Sequence 22, Appl	200	5	17.9	968	1	US-08-426-236-2	Sequence 2, Appli
128	5	17.9	456	1	US-08-464-164-2	Sequence 2, Appli	201	5	17.9	997	1	US-08-324-977-50	Sequence 50, Appl
129	5	17.9	456	1	US-08-338-057-2	Sequence 2, Appli	202	5	17.9	997	2	US-08-384-616-50	Sequence 50, Appl
130	5	17.9	456	2	US-08-668-416-2	Sequence 2, Appli	203	5	17.9	997	2	US-08-904-886A-50	Sequence 50, Appl
131	5	17.9	478	2	US-08-873-093-1	Sequence 1, Appli	204	5	17.9	1005	2	US-09-315-850-50	Sequence 50, Appl
132	5	17.9	478	2	US-08-873-093-4	Sequence 4, Appli	205	5	17.9	1042	4	US-09-387-695-2	Sequence 2, Appli
133	5	17.9	480	2	US-08-724-394A-11	Sequence 11, Appl	206	5	17.9	1274	4	US-09-095-443-2	Sequence 2, Appli
134	5	17.9	496	4	US-09-292-768-64	Sequence 64, Appl	207	5	17.9	1274	4	US-09-095-443-2	Sequence 2, Appli
135	5	17.9	523	1	US-08-073-383-2	Sequence 2, Appli	208	5	17.9	1296	1	US-08-480-604A-28	Sequence 28, Appl
136	5	17.9	523	1	US-08-428-415-2	Sequence 2, Appli	209	5	17.9	1296	4	US-08-405-496A-28	Sequence 28, Appl
137	5	17.9	523	1	US-08-379-685-2	Sequence 2, Appli	210	5	17.9	1296	4	US-08-915-136-28	Sequence 28, Appl
138	5	17.9	523	2	US-08-854-029-2	Sequence 2, Appli	211	5	17.9	1420	2	US-08-540-804-14	Sequence 14, Appl
139	5	17.9	523	4	US-08-428-762-2	Sequence 2, Appli	212	5	17.9	1420	2	US-08-218-265-14	Sequence 14, Appl
140	5	17.9	523	5	PCT-US94-06365-2	Sequence 2, Appli	213	5	17.9	1420	3	US-08-521-872-14	Sequence 14, Appl
141	5	17.9	540	4	US-09-433-994-2	Sequence 2, Appli	214	5	17.9	1420	4	US-08-590-399-14	Sequence 14, Appl
142	5	17.9	542	4	US-09-380-420C-17	Sequence 17, Appl	215	5	17.9	1463	1	US-08-157-005-3	Sequence 3, Appli
143	5	17.9	546	2	US-08-492-027A-8	Sequence 8, Appli	216	5	17.9	1463	4	US-08-747-863-3	Sequence 3, Appli
144	5	17.9	560	1	US-07-683-957B-1	Sequence 1, Appli	217	5	17.9	1528	1	US-08-326-117B-2	Sequence 2, Appli
145	5	17.9	603	4	US-09-042-785A-23	Sequence 23, Appl	218	5	17.9	1528	3	US-08-982-129-2	Sequence 2, Appli
146	5	17.9	609	1	US-07-798-776-2	Sequence 2, Appli	219	5	17.9	1544	4	US-09-413-814-46	Sequence 46, Appl
147	5	17.9	609	3	US-08-251-288A-2	Sequence 2, Appli	220	5	17.9	1663	2	US-08-793-126-1	Sequence 1, Appli
148	5	17.9	609	3	US-09-298-819A-2	Sequence 2, Appli	221	5	17.9	1663	4	US-09-132-271-1	Sequence 1, Appli
149	5	17.9	620	4	US-09-269-731-6	Sequence 6, Appli	222	5	17.9	1663	4	US-09-142-334-22	Sequence 22, Appl
150	5	17.9	630	2	US-08-797-366-3	Sequence 3, Appli	223	5	17.9	1864	2	US-08-804-227C-3	Sequence 3, Appli
151	5	17.9	630	2	US-08-956-268-3	Sequence 3, Appli	224	5	17.9	1898	1	US-08-056-200-94	Sequence 94, Appl
152	5	17.9	650	4	US-09-232-191-29	Sequence 29, Appl	225	5	17.9	1898	2	US-08-800-644-94	Sequence 94, Appl
153	5	17.9	650	4	US-09-232-200-79	Sequence 79, Appl	226	5	17.9	2101	1	US-08-466-390-4	Sequence 4, Appli
154	5	17.9	650	4	US-09-232-200-97	Sequence 97, Appl	227	5	17.9	2101	1	US-08-470-950-4	Sequence 4, Appli
155	5	17.9	650	4	US-09-232-197-79	Sequence 79, Appl	228	5	17.9	2101	1	US-08-467-781-4	Sequence 4, Appli
156	5	17.9	650	4	US-09-232-197-97	Sequence 97, Appl	229	5	17.9	2101	1	US-08-195-487-4	Sequence 4, Appli
157	5	17.9	650	4	US-09-232-201-79	Sequence 79, Appl	230	5	17.9	2101	2	US-08-483-924-4	Sequence 4, Appli
158	5	17.9	650	4	US-09-232-201-97	Sequence 97, Appl	231	5	17.9	2101	4	US-09-452-294-1	Sequence 1, Appli
159	5	17.9	651	2	US-08-492-027A-1	Sequence 1, Appli	232	5	17.9	2101	5	PCT-US93-06160-4	Sequence 4, Appli
160	5	17.9	655	2	US-08-492-027A-6	Sequence 6, Appli	233	5	17.9	2620	2	US-08-324-977-32	Sequence 32, Appl
161	5	17.9	676	4	US-09-693-147-6	Sequence 6, Appli	234	5	17.9	2620	2	US-08-384-616-32	Sequence 32, Appl
162	5	17.9	682	4	US-07-757-342D-6	Sequence 6, Appli	235	5	17.9	2620	2	US-08-904-686A-32	Sequence 32, Appl
163	5	17.9	728	4	US-08-981-392-2	Sequence 2, Appli	236	5	17.9	2620	4	US-09-315-850-32	Sequence 32, Appl
164	5	17.9	729	3	US-08-872-855-8	Sequence 8, Appli	237	5	17.9	2621	1	US-08-324-977-36	Sequence 36, Appl
165	5	17.9	766	2	US-08-846-762-5	Sequence 5, Appli	238	5	17.9	2621	2	US-08-384-616-36	Sequence 36, Appl
166	5	17.9	771	3	US-09-121-964-9	Sequence 9, Appli	239	5	17.9	2621	2	US-08-904-686A-36	Sequence 36, Appl
167	5	17.9	803	1	US-08-158-232-10	Sequence 10, Appl	240	5	17.9	2621	4	US-09-315-850-36	Sequence 36, Appl
168	5	17.9	803	1	US-08-304-626-10	Sequence 10, Appl	241	5	17.9	3010	1	US-08-324-977-2	Sequence 2, Appli
169	5	17.9	803	1	US-08-316-301A-12	Sequence 12, Appl	242	5	17.9	3010	1	US-08-324-977-14	Sequence 14, Appl
170	5	17.9	803	2	US-08-611-928-10	Sequence 10, Appl	243	5	17.9	3010	2	US-08-384-616-2	Sequence 2, Appli
171	5	17.9	803	3	US-09-173-891-10	Sequence 10, Appl	244	5	17.9	3010	2	US-08-384-616-14	Sequence 14, Appl
172	5	17.9	803	4	US-09-076-137-12	Sequence 12, Appl	245	5	17.9	3010	2	US-08-904-686A-2	Sequence 2, Appli
173	5	17.9	803	5	PCT-US92-03624-12	Sequence 12, Appl	246	5	17.9	3010	2	US-08-904-686A-14	Sequence 14, Appl

247	5	17.9	3010	4	US-09-315-850-2	Sequence 2, Appli	320	4	14.3	14	2	US-08-253-678A-14	Sequence 14, Appl
248	5	17.9	3010	4	US-09-315-850-14	Sequence 14, Appl	321	4	14.3	14	3	US-08-582-134B-14	Sequence 14, Appl
249	5	17.9	3079	5	PCT-US94-00198-4	Sequence 4, Appli	322	4	14.3	14	3	US-08-170-299-15	Sequence 15, Appl
250	5	17.9	3170	2	US-07-642-734C-5	Sequence 5, Appli	323	4	14.3	14	3	US-08-467-791-13	Sequence 13, Appl
251	5	17.9	3170	3	US-08-433-009A-5	Sequence 6, Appli	324	4	14.3	14	6	544158-4	Patent No. 544158
252	5	17.9	5588	4	US-09-036-987A-6	Sequence 6, Appli	325	4	14.3	15	1	US-08-147-000B-18	Sequence 18, Appl
253	5	17.9	5588	4	US-09-370-700-6	Sequence 6, Appli	326	4	14.3	16	1	US-07-854-629-5	Sequence 5, Appli
254	4	14.3	5	2	US-08-350-260A-327	Sequence 327, App	327	4	14.3	16	1	US-08-474-008-5	Sequence 1, Appli
255	4	14.3	5	4	US-09-236-160-9	Sequence 9, Appli	328	4	14.3	16	2	US-08-908-526-1	Sequence 1, Appli
256	4	14.3	7	2	US-08-680-326-94	Sequence 94, Appli	329	4	14.3	17	2	US-08-480-434-25	Sequence 25, Appl
257	4	14.3	8	4	US-08-444-818-433	Sequence 433, App	330	4	14.3	17	2	US-08-480-434-33	Sequence 33, Appl
258	4	14.3	9	2	US-08-908-526-6	Sequence 6, Appli	331	4	14.3	17	2	US-08-712-212-15	Sequence 15, Appl
259	4	14.3	9	4	US-09-139-802-108	Sequence 108, App	332	4	14.3	17	2	US-08-053-451B-25	Sequence 25, Appl
260	4	14.3	9	4	US-08-793-701-29	Sequence 29, Appl	333	4	14.3	17	2	US-08-053-451B-33	Sequence 33, Appl
261	4	14.3	9	4	US-08-793-701-31	Sequence 31, Appl	334	4	14.3	17	4	US-09-025-596-93	Sequence 93, Appl
262	4	14.3	9	4	US-08-461-697-401	Sequence 401, App	335	4	14.3	17	5	PCT-US95-05160-15	Sequence 15, Appl
263	4	14.3	10	2	US-08-406-330-17	Sequence 17, Appl	336	4	14.3	17	5	PCT-US95-08621-25	Sequence 25, Appl
264	4	14.3	10	2	US-08-556-597-17	Sequence 17, Appl	337	4	14.3	18	1	US-08-170-596-1	Sequence 1, Appli
265	4	14.3	10	4	US-09-113-312-6	Sequence 6, Appli	338	4	14.3	18	1	US-08-383-753-60	Sequence 60, Appl
266	4	14.3	10	6	5428134-1	Patent No. 5428134	339	4	14.3	18	2	US-08-586-772-60	Sequence 60, Appl
267	4	14.3	10	6	5428134-10	Patent No. 5428134	340	4	14.3	18	2	US-08-959-512-60	Sequence 60, Appl
268	4	14.3	10	6	5436320-1	Patent No. 5436320	341	4	14.3	18	4	US-09-512-983-60	Sequence 60, Appl
269	4	14.3	10	6	5436320-7	Patent No. 5436320	342	4	14.3	18	4	US-09-227-357-510	Sequence 510, App
270	4	14.3	11	1	US-07-868-353A-7	Sequence 7, Appli	343	4	14.3	19	1	US-07-805-437-4	Sequence 4, Appli
271	4	14.3	11	2	US-08-407-804-7	Sequence 7, Appli	344	4	14.3	19	1	US-07-805-437-5	Sequence 5, Appli
272	4	14.3	11	2	US-08-360-606B-1	Sequence 1, Appli	345	4	14.3	19	2	US-08-635-062-5	Sequence 5, Appli
273	4	14.3	11	3	US-09-124-807-7	Sequence 7, Appli	346	4	14.3	19	4	US-09-330-740A-6	Sequence 6, Appli
274	4	14.3	11	4	US-08-842-306B-33	Sequence 33, Appl	347	4	14.3	20	1	US-08-250-789A-176	Sequence 176, App
275	4	14.3	11	4	US-08-842-306B-34	Sequence 34, Appl	348	4	14.3	20	2	US-08-419-061A-8	Sequence 8, Appli
276	4	14.3	11	4	US-08-838-973B-33	Sequence 33, Appl	349	4	14.3	20	2	US-08-447-430A-25	Sequence 25, Appl
277	4	14.3	11	4	US-08-838-973B-34	Sequence 34, Appl	350	4	14.3	20	2	US-08-107-676-16	Sequence 16, Appl
278	4	14.3	12	1	US-08-486-135-9	Sequence 9, Appli	351	4	14.3	20	2	US-08-053-451B-116	Sequence 116, App
279	4	14.3	12	1	US-08-486-135-16	Sequence 16, Appl	352	4	14.3	20	2	US-08-485-647A-8	Sequence 8, Appli
280	4	14.3	12	1	US-08-468-975-6	Sequence 6, Appli	353	4	14.3	20	4	US-09-133-341-4	Sequence 4, Appli
281	4	14.3	12	1	US-08-468-975-12	Sequence 12, Appl	354	4	14.3	20	5	PCT-US94-10257A-87	Sequence 87, Appl
282	4	14.3	12	1	US-08-470-152-9	Sequence 9, Appli	355	4	14.3	20	5	PCT-US94-10257A-88	Sequence 88, Appl
283	4	14.3	12	1	US-08-470-152-16	Sequence 16, Appl	356	4	14.3	21	1	US-08-486-135-10	Sequence 10, Appl
284	4	14.3	12	2	US-08-468-964B-7	Sequence 7, Appli	357	4	14.3	21	1	US-08-468-543-25	Sequence 25, Appl
285	4	14.3	12	2	US-07-871-282A-7	Sequence 7, Appli	358	4	14.3	21	1	US-08-468-975-7	Sequence 7, Appli
286	4	14.3	12	2	US-08-253-678A-7	Sequence 7, Appli	359	4	14.3	21	1	US-08-470-152-10	Sequence 10, Appl
287	4	14.3	12	2	US-08-253-678A-13	Sequence 13, Appl	360	4	14.3	21	2	US-08-468-964B-8	Sequence 8, Appli
288	4	14.3	12	3	US-08-582-134B-7	Sequence 7, Appli	361	4	14.3	21	2	US-08-845-926-35	Sequence 35, Appl
289	4	14.3	12	3	US-08-582-134B-13	Sequence 13, Appl	362	4	14.3	21	2	US-08-469-692-25	Sequence 25, Appl
290	4	14.3	12	3	US-08-170-299-7	Sequence 7, Appli	363	4	14.3	21	2	US-07-871-282A-8	Sequence 8, Appli
291	4	14.3	12	3	US-08-170-299-14	Sequence 14, Appl	364	4	14.3	21	2	US-08-398-046-25	Sequence 25, Appl
292	4	14.3	12	3	US-08-467-791-6	Sequence 6, Appli	365	4	14.3	21	2	US-08-466-860-74	Sequence 74, Appl
293	4	14.3	12	3	US-08-467-791-12	Sequence 12, Appl	366	4	14.3	21	2	US-08-253-678A-8	Sequence 8, Appli
294	4	14.3	13	2	US-08-292-968-40	Sequence 40, Appl	367	4	14.3	21	3	US-08-582-134B-8	Sequence 8, Appli
295	4	14.3	13	2	US-08-467-974-40	Sequence 40, Appl	368	4	14.3	21	3	US-08-170-299-8	Sequence 8, Appli
296	4	14.3	13	2	US-08-467-536-40	Sequence 40, Appl	369	4	14.3	21	3	US-08-472-040A-74	Sequence 74, Appl
297	4	14.3	13	3	US-08-467-976-40	Sequence 40, Appl	370	4	14.3	21	3	US-08-467-791-7	Sequence 7, Appli
298	4	14.3	13	3	US-08-582-740-64	Sequence 64, Appl	371	4	14.3	21	4	US-08-276-776-74	Sequence 74, Appl
299	4	14.3	13	3	US-09-188-579-17	Sequence 17, Appl	372	4	14.3	21	4	US-08-471-209-74	Sequence 74, Appl
300	4	14.3	13	4	US-09-082-514-40	Sequence 40, Appl	373	4	14.3	22	1	US-07-672-483-13	Sequence 13, Appl
301	4	14.3	13	4	US-08-602-999A-88	Sequence 88, Appl	374	4	14.3	22	1	US-07-672-483-14	Sequence 14, Appl
302	4	14.3	13	4	US-08-602-999A-89	Sequence 89, Appl	375	4	14.3	22	1	US-08-419-009-10	Sequence 10, Appl
303	4	14.3	13	4	US-08-664-962B-20	Sequence 20, Appl	376	4	14.3	22	2	US-08-360-606B-12	Sequence 12, Appl
304	4	14.3	13	4	US-09-315-444-17	Sequence 17, Appl	377	4	14.3	22	2	US-08-969-721-6	Sequence 6, Appli
305	4	14.3	13	4	US-09-311-743-20	Sequence 20, Appl	378	4	14.3	23	1	US-08-192-310-2	Sequence 2, Appli
306	4	14.3	13	4	US-09-109-879-64	Sequence 64, Appl	379	4	14.3	23	1	US-08-231-730A-8	Sequence 8, Appli
307	4	14.3	13	4	US-08-278-865-88	Sequence 88, Appl	380	4	14.3	23	1	US-08-457-798-8	Sequence 8, Appli
308	4	14.3	13	4	US-08-278-865-89	Sequence 89, Appl	381	4	14.3	23	2	US-08-457-171-8	Sequence 8, Appli
309	4	14.3	13	4	US-08-278-865-90	Sequence 90, Appl	382	4	14.3	23	2	US-08-712-212-4	Sequence 4, Appli
310	4	14.3	13	5	PCT-US94-10257A-64	Sequence 64, Appl	383	4	14.3	23	2	US-08-505-486-8	Sequence 8, Appli
311	4	14.3	13	5	PCT-US94-10257A-65	Sequence 65, Appl	384	4	14.3	23	2	US-08-475-328-8	Sequence 8, Appli
312	4	14.3	14	1	US-08-486-135-17	Sequence 17, Appl	385	4	14.3	23	3	US-08-689-489C-8	Sequence 8, Appli
313	4	14.3	14	1	US-08-468-975-13	Sequence 13, Appl	386	4	14.3	23	3	US-08-801-028-8	Sequence 8, Appli
314	4	14.3	14	1	US-08-470-152-17	Sequence 17, Appl	387	4	14.3	23	3	US-09-340-154-8	Sequence 8, Appli
315	4	14.3	14	2	US-08-460-309-23	Sequence 23, Appl	388	4	14.3	23	3	US-08-851-843A-208	Sequence 208, App
316	4	14.3	14	2	US-08-125-077-23	Sequence 23, Appl	389	4	14.3	23	4	US-08-974-549A-327	Sequence 327, App
317	4	14.3	14	2	US-08-908-526-5	Sequence 5, Appli	390	4	14.3	23	4	US-09-232-802A-8	Sequence 8, Appli
318	4	14.3	14	2	US-08-468-964B-13	Sequence 13, Appl	391	4	14.3	23	4	US-08-854-050-208	Sequence 208, App
319	4	14.3	14	2	US-07-871-282A-13	Sequence 13, Appl	392	4	14.3	23	4	US-09-430-323-208	Sequence 208, App

393	4	14.3	23	5	PCT-US94-06176-8	Sequence 8, Appli	456	4	14.3	28	1	US-08-786-748A-19	Sequence 19, Appli
394	4	14.3	23	5	PCT-US95-04335-8	Sequence 8, Appli	467	4	14.3	28	1	US-08-786-748A-160	Sequence 160, App
395	4	14.3	23	5	PCT-US95-04718-8	Sequence 8, Appli	468	4	14.3	28	2	US-08-932-682-19	Sequence 19, Appli
396	4	14.3	23	5	PCT-US95-05160-4	Sequence 4, Appli	469	4	14.3	28	2	US-08-932-682-160	Sequence 160, App
397	4	14.3	23	5	PCT-US95-09338-8	Sequence 8, Appli	470	4	14.3	28	4	US-09-253-396A-221	Sequence 221, App
398	4	14.3	23	5	PCT-US95-09339-8	Sequence 8, Appli	471	4	14.3	28	4	US-09-082-279B-52	Sequence 52, Appli
399	4	14.3	24	1	US-07-976-338-24	Sequence 24, Appli	472	4	14.3	28	4	US-09-082-279B-53	Sequence 53, Appli
400	4	14.3	24	1	US-08-484-635-83	Sequence 83, Appli	473	4	14.3	28	4	US-09-082-279B-1279	Sequence 1279, Ap
401	4	14.3	24	2	US-08-484-631-83	Sequence 83, Appli	474	4	14.3	28	4	US-09-082-279B-1282	Sequence 1282, Ap
402	4	14.3	24	2	US-08-481-793-26	Sequence 26, Appli	475	4	14.3	28	4	US-09-082-279B-1314	Sequence 1314, Ap
403	4	14.3	24	2	US-08-481-793-28	Sequence 28, Appli	476	4	14.3	28	4	US-09-082-279B-1317	Sequence 1317, Ap
404	4	14.3	24	2	US-08-481-793-29	Sequence 29, Appli	477	4	14.3	28	4	US-09-105-350-32	Sequence 32, Appli
405	4	14.3	24	2	US-08-354-326-28	Sequence 26, Appli	478	4	14.3	28	4	US-09-315-304B-52	Sequence 52, Appli
406	4	14.3	24	2	US-08-354-326-36	Sequence 28, Appli	479	4	14.3	28	4	US-09-315-304B-53	Sequence 53, Appli
407	4	14.3	24	2	US-08-354-326-29	Sequence 29, Appli	480	4	14.3	28	4	US-09-315-304B-1279	Sequence 1279, Ap
408	4	14.3	24	2	US-08-827-570-83	Sequence 83, Appli	481	4	14.3	28	4	US-09-315-304B-1282	Sequence 1282, Ap
409	4	14.3	24	2	US-08-827-376-2	Sequence 2, Appli	482	4	14.3	28	4	US-09-315-304B-1314	Sequence 1314, Ap
410	4	14.3	24	5	PCT-US95-07068-26	Sequence 26, Appli	483	4	14.3	28	4	US-09-315-304B-1317	Sequence 1317, Ap
411	4	14.3	24	5	PCT-US95-07068-28	Sequence 28, Appli	484	4	14.3	29	1	US-08-520-599-3	Sequence 3, Appli
412	4	14.3	24	5	PCT-US95-07068-29	Sequence 29, Appli	485	4	14.3	29	2	US-08-310-912A-152	Sequence 152, App
413	4	14.3	25	1	US-07-976-338-15	Sequence 15, Appli	486	4	14.3	29	4	US-08-905-223-423	Sequence 423, App
414	4	14.3	25	1	US-07-976-338-21	Sequence 21, Appli	487	4	14.3	29	4	US-09-301-085-152	Sequence 152, App
415	4	14.3	25	1	US-07-956-700B-69	Sequence 69, Appli	488	4	14.3	29	4	US-08-860-656B-15	Sequence 15, Appli
416	4	14.3	25	1	US-08-476-537-69	Sequence 69, Appli	489	4	14.3	29	5	PCT-US95-04589-152	Sequence 152, App
417	4	14.3	25	1	US-08-485-607-69	Sequence 69, Appli	490	4	14.3	29	6	5221820-7	Patent No. 5221820
418	4	14.3	25	2	US-08-473-475A-8	Sequence 8, Appli	491	4	14.3	30	1	US-08-231-730A-11	Sequence 11, Appli
419	4	14.3	25	2	US-08-475-879-69	Sequence 69, Appli	492	4	14.3	30	1	US-08-457-798-11	Sequence 11, Appli
420	4	14.3	25	5	PCT-US92-07813-10	Sequence 10, Appli	493	4	14.3	30	1	US-08-457-171-11	Sequence 11, Appli
421	4	14.3	25	5	PCT-US94-04361-36	Sequence 36, Appli	494	4	14.3	30	2	US-08-473-475A-2	Sequence 2, Appli
422	4	14.3	26	1	US-07-626-589-6	Sequence 6, Appli	495	4	14.3	30	2	US-08-419-061A-6	Sequence 6, Appli
423	4	14.3	26	1	US-08-323-444A-11	Sequence 11, Appli	496	4	14.3	30	2	US-08-475-328-11	Sequence 11, Appli
424	4	14.3	26	1	US-08-231-730A-12	Sequence 12, Appli	497	4	14.3	30	2	US-08-485-647A-6	Sequence 6, Appli
425	4	14.3	26	1	US-08-236-410-6	Sequence 6, Appli	498	4	14.3	30	3	US-08-689-489C-11	Sequence 11, Appli
426	4	14.3	26	1	US-08-213-124-7	Sequence 7, Appli	499	4	14.3	30	3	US-09-164-021-7	Sequence 7, Appli
427	4	14.3	26	1	US-08-457-798-12	Sequence 12, Appli	500	4	14.3	30	4	US-09-232-802A-11	Sequence 11, Appli
428	4	14.3	26	1	US-08-457-798-12	Sequence 12, Appli	501	4	14.3	30	5	PCT-US94-06176-11	Sequence 11, Appli
429	4	14.3	26	1	US-08-457-171-12	Sequence 12, Appli	502	4	14.3	30	5	PCT-US95-04335-11	Sequence 11, Appli
430	4	14.3	26	1	US-08-455-421-6	Sequence 1, Appli	503	4	14.3	30	5	PCT-US95-04718-11	Sequence 11, Appli
431	4	14.3	26	2	US-08-476-062A-1	Sequence 1, Appli	504	4	14.3	31	1	US-08-190-802A-209	Sequence 209, App
432	4	14.3	26	2	US-08-620-151-93	Sequence 93, Appli	505	4	14.3	31	1	US-08-190-802A-219	Sequence 219, App
433	4	14.3	26	2	US-08-505-486-12	Sequence 12, Appli	506	4	14.3	31	2	US-08-505-486-11	Sequence 11, Appli
434	4	14.3	26	2	US-08-475-328-12	Sequence 12, Appli	507	4	14.3	31	3	US-08-801-028-11	Sequence 11, Appli
435	4	14.3	26	3	US-08-689-489C-12	Sequence 12, Appli	508	4	14.3	31	3	US-09-340-154-11	Sequence 11, Appli
436	4	14.3	26	3	US-08-801-028-12	Sequence 12, Appli	509	4	14.3	31	4	US-08-602-999A-56	Sequence 56, Appli
437	4	14.3	26	3	US-09-340-154-12	Sequence 12, Appli	510	4	14.3	31	4	US-08-602-999A-57	Sequence 57, Appli
438	4	14.3	26	3	US-09-082-306-12	Sequence 12, Appli	511	4	14.3	31	4	US-08-477-346-209	Sequence 209, App
439	4	14.3	26	4	US-09-232-802A-12	Sequence 12, Appli	512	4	14.3	31	4	US-08-477-346-219	Sequence 219, App
440	4	14.3	26	4	US-08-905-223-430	Sequence 430, App	513	4	14.3	31	4	US-08-278-865-56	Sequence 56, Appli
441	4	14.3	26	4	US-05-208-684-5	Sequence 5, Appli	514	4	14.3	31	4	US-08-278-865-57	Sequence 57, Appli
442	4	14.3	26	5	PCT-US94-06176-12	Sequence 35, Appli	515	4	14.3	31	4	US-08-473-089-209	Sequence 209, App
443	4	14.3	26	5	PCT-US95-04335-12	Sequence 12, Appli	516	4	14.3	31	4	US-08-473-089-219	Sequence 219, App
444	4	14.3	26	5	PCT-US95-04718-12	Sequence 12, Appli	517	4	14.3	31	5	PCT-US95-09338-11	Sequence 11, Appli
445	4	14.3	26	5	PCT-US95-09338-12	Sequence 12, Appli	518	4	14.3	31	5	PCT-US95-09339-11	Sequence 11, Appli
446	4	14.3	26	5	PCT-US95-09339-12	Sequence 12, Appli	519	4	14.3	32	1	US-08-190-802A-181	Sequence 181, App
447	4	14.3	26	5	PCT-US95-09339-11	Sequence 1, Appli	520	4	14.3	32	1	US-07-988-925-9	Sequence 9, Appli
448	4	14.3	27	1	US-08-231-730A-7	Sequence 7, Appli	521	4	14.3	32	1	US-08-471-780C-82	Sequence 82, Appli
449	4	14.3	27	1	US-08-786-730A-7	Sequence 7, Appli	522	4	14.3	32	1	US-07-977-696C-79	Sequence 79, Appli
450	4	14.3	27	1	US-08-457-798-7	Sequence 82, Appli	523	4	14.3	32	1	US-08-467-282B-82	Sequence 82, Appli
451	4	14.3	27	1	US-08-457-171-7	Sequence 7, Appli	524	4	14.3	32	1	US-08-129-930B-79	Sequence 79, Appli
452	4	14.3	27	1	US-08-457-171-7	Sequence 7, Appli	525	4	14.3	32	2	US-08-292-870-3	Sequence 3, Appli
453	4	14.3	27	2	US-08-505-486-7	Sequence 82, Appli	526	4	14.3	32	2	US-08-471-282A-82	Sequence 82, Appli
454	4	14.3	27	2	US-08-475-328-7	Sequence 7, Appli	527	4	14.3	32	2	US-08-466-710C-82	Sequence 82, Appli
455	4	14.3	27	3	US-08-689-489C-7	Sequence 7, Appli	528	4	14.3	32	2	US-08-362-780-9	Sequence 9, Appli
456	4	14.3	27	3	US-08-801-028-7	Sequence 7, Appli	529	4	14.3	32	3	US-08-470-139-21	Sequence 21, Appli
457	4	14.3	27	3	US-08-801-028-7	Sequence 7, Appli	530	4	14.3	32	3	US-08-938-548B-7	Sequence 7, Appli
458	4	14.3	27	3	US-09-340-154-7	Sequence 7, Appli	531	4	14.3	32	3	US-08-468-739C-82	Sequence 82, Appli
459	4	14.3	27	5	PCT-US94-06176-7	Sequence 7, Appli	532	4	14.3	32	4	US-08-477-346-181	Sequence 181, App
460	4	14.3	27	5	PCT-US95-04335-7	Sequence 7, Appli	533	4	14.3	32	4	US-08-939-093A-7	Sequence 7, Appli
461	4	14.3	27	5	PCT-US95-04718-7	Sequence 7, Appli	534	4	14.3	32	4	US-08-976-288A-79	Sequence 79, Appli
462	4	14.3	27	5	PCT-US95-09338-7	Sequence 7, Appli	535	4	14.3	32	4	US-09-347-061-21	Sequence 21, Appli
463	4	14.3	27	5	PCT-US95-09339-7	Sequence 7, Appli	536	4	14.3	32	4	US-08-473-089-181	Sequence 181, App
464	4	14.3	27	5	PCT-US95-09339-7	Sequence 7, Appli	537	4	14.3	32	4	US-09-425-638A-108	Sequence 108, App
465	4	14.3	28	1	US-08-279-058B-1	Sequence 1, Appli	538	4	14.3	32	4	US-09-425-638A-109	Sequence 109, App

539	4	14.3	32	4	US-09-425-638A-110	Sequence 110, App	612	4	14.3	35	3	US-08-486-099-69	Sequence 69, Appl
540	4	14.3	32	4	US-09-425-638A-111	Sequence 111, App	613	4	14.3	35	3	US-08-486-099-70	Sequence 70, Appl
541	4	14.3	32	4	US-09-425-638A-112	Sequence 112, App	614	4	14.3	35	3	US-08-486-099-71	Sequence 71, Appl
542	4	14.3	32	4	US-09-425-638A-113	Sequence 113, App	615	4	14.3	35	3	US-08-486-099-72	Sequence 72, Appl
543	4	14.3	32	4	US-09-425-638A-114	Sequence 114, App	616	4	14.3	35	3	US-08-273-685-22	Sequence 22, Appl
544	4	14.3	32	4	US-09-425-638A-115	Sequence 115, App	617	4	14.3	35	3	US-08-360-107A-63	Sequence 63, Appl
545	4	14.3	32	4	US-09-425-638A-116	Sequence 116, App	618	4	14.3	35	3	US-08-360-107A-64	Sequence 64, Appl
546	4	14.3	32	4	US-09-425-638A-117	Sequence 117, App	619	4	14.3	35	3	US-08-360-107A-65	Sequence 65, Appl
547	4	14.3	32	4	US-09-425-638A-118	Sequence 118, App	620	4	14.3	35	3	US-08-360-107A-66	Sequence 66, Appl
548	4	14.3	32	4	US-09-425-638A-119	Sequence 119, App	621	4	14.3	35	3	US-08-360-107A-67	Sequence 67, Appl
549	4	14.3	32	4	US-09-425-638A-120	Sequence 120, App	622	4	14.3	35	3	US-08-360-107A-68	Sequence 68, Appl
550	4	14.3	32	4	US-09-425-638A-121	Sequence 121, App	623	4	14.3	35	3	US-08-360-107A-69	Sequence 69, Appl
551	4	14.3	32	4	US-09-425-638A-122	Sequence 122, App	624	4	14.3	35	3	US-08-360-107A-70	Sequence 70, Appl
552	4	14.3	32	4	US-09-425-638A-123	Sequence 123, App	625	4	14.3	35	3	US-08-360-107A-71	Sequence 71, Appl
553	4	14.3	32	4	US-09-425-638A-124	Sequence 124, App	626	4	14.3	35	3	US-08-360-107A-72	Sequence 72, Appl
554	4	14.3	32	4	US-09-543-004-108	Sequence 108, App	627	4	14.3	35	3	US-08-360-107A-119	Sequence 139, App
555	4	14.3	32	4	US-09-543-004-109	Sequence 109, App	628	4	14.3	35	3	US-08-360-107A-140	Sequence 140, App
556	4	14.3	32	4	US-09-543-004-110	Sequence 110, App	629	4	14.3	35	3	US-08-360-107A-141	Sequence 141, App
557	4	14.3	32	4	US-09-543-004-111	Sequence 111, App	630	4	14.3	35	3	US-08-360-107A-142	Sequence 142, App
558	4	14.3	32	4	US-09-543-004-112	Sequence 112, App	631	4	14.3	35	3	US-08-360-107A-143	Sequence 143, App
559	4	14.3	32	4	US-09-543-004-113	Sequence 113, App	632	4	14.3	35	3	US-08-360-107A-144	Sequence 144, App
560	4	14.3	32	4	US-09-543-004-114	Sequence 114, App	633	4	14.3	35	3	US-08-360-107A-145	Sequence 145, App
561	4	14.3	32	4	US-09-543-004-115	Sequence 115, App	634	4	14.3	35	3	US-08-360-107A-146	Sequence 146, App
562	4	14.3	32	4	US-09-543-004-116	Sequence 116, App	635	4	14.3	35	3	US-08-360-107A-147	Sequence 147, App
563	4	14.3	32	4	US-09-543-004-117	Sequence 117, App	636	4	14.3	35	3	US-08-360-107A-148	Sequence 148, App
564	4	14.3	32	4	US-09-543-004-118	Sequence 118, App	637	4	14.3	35	3	US-08-360-107A-149	Sequence 149, App
565	4	14.3	32	4	US-09-543-004-119	Sequence 119, App	638	4	14.3	35	3	US-08-484-223B-63	Sequence 63, Appl
566	4	14.3	32	4	US-09-543-004-120	Sequence 120, App	639	4	14.3	35	3	US-08-484-223B-64	Sequence 64, Appl
567	4	14.3	32	4	US-09-543-004-121	Sequence 121, App	640	4	14.3	35	3	US-08-484-223B-65	Sequence 65, Appl
568	4	14.3	32	4	US-09-543-004-122	Sequence 122, App	641	4	14.3	35	3	US-08-484-223B-66	Sequence 66, Appl
569	4	14.3	32	4	US-09-543-004-123	Sequence 123, App	642	4	14.3	35	3	US-08-484-223B-67	Sequence 67, Appl
570	4	14.3	32	4	US-09-543-004-124	Sequence 124, App	643	4	14.3	35	3	US-08-484-223B-68	Sequence 68, Appl
571	4	14.3	33	1	US-08-086-428B-137	Sequence 137, App	644	4	14.3	35	3	US-08-484-223B-69	Sequence 69, Appl
572	4	14.3	33	1	US-08-086-428B-138	Sequence 138, App	645	4	14.3	35	3	US-08-484-223B-70	Sequence 70, Appl
573	4	14.3	33	1	US-08-520-599-2	Sequence 2, Appl1	646	4	14.3	35	3	US-08-484-223B-71	Sequence 71, Appl
574	4	14.3	33	2	US-08-468-570-137	Sequence 137, App	647	4	14.3	35	3	US-08-484-223B-72	Sequence 72, Appl
575	4	14.3	33	2	US-08-468-570-138	Sequence 138, App	648	4	14.3	35	3	US-08-919-597-63	Sequence 63, Appl
576	4	14.3	33	2	US-08-290-665A-241	Sequence 241, App	649	4	14.3	35	3	US-08-919-597-64	Sequence 64, Appl
577	4	14.3	33	2	US-08-290-665A-242	Sequence 242, App	650	4	14.3	35	3	US-08-919-597-65	Sequence 65, Appl
578	4	14.3	33	2	US-08-690-011A-38	Sequence 38, Appl	651	4	14.3	35	3	US-08-919-597-66	Sequence 66, Appl
579	4	14.3	33	2	US-08-511-872-4	Sequence 4, Appl1	652	4	14.3	35	3	US-08-919-597-67	Sequence 67, Appl
580	4	14.3	33	2	US-08-511-872-5	Sequence 5, Appl1	653	4	14.3	35	3	US-08-919-597-68	Sequence 68, Appl
581	4	14.3	33	2	US-08-511-872-6	Sequence 6, Appl1	654	4	14.3	35	3	US-08-919-597-69	Sequence 69, Appl
582	4	14.3	33	2	US-08-511-872-7	Sequence 7, Appl1	655	4	14.3	35	3	US-08-919-597-70	Sequence 70, Appl
583	4	14.3	33	2	US-08-511-872-8	Sequence 8, Appl1	656	4	14.3	35	3	US-08-919-597-71	Sequence 71, Appl
584	4	14.3	33	2	US-08-511-872-9	Sequence 9, Appl1	657	4	14.3	35	3	US-08-919-597-72	Sequence 72, Appl
585	4	14.3	33	2	US-08-511-872-10	Sequence 10, Appl	658	4	14.3	35	3	US-08-475-668A-63	Sequence 63, Appl
586	4	14.3	33	2	US-08-511-872-11	Sequence 11, Appl	659	4	14.3	35	3	US-08-475-668A-64	Sequence 64, Appl
587	4	14.3	33	2	US-08-310-912A-34	Sequence 34, Appl	660	4	14.3	35	3	US-08-475-668A-65	Sequence 65, Appl
588	4	14.3	33	4	US-08-841-089-35	Sequence 35, Appl	661	4	14.3	35	3	US-08-475-668A-66	Sequence 66, Appl
589	4	14.3	33	4	US-08-082-279B-399	Sequence 399, App	662	4	14.3	35	3	US-08-475-668A-67	Sequence 67, Appl
590	4	14.3	33	4	US-08-981-392-63	Sequence 63, Appl	663	4	14.3	35	3	US-08-475-668A-68	Sequence 68, Appl
591	4	14.3	33	4	US-09-301-085-34	Sequence 34, Appl	664	4	14.3	35	3	US-08-475-668A-69	Sequence 69, Appl
592	4	14.3	33	4	US-09-315-304B-399	Sequence 399, App	665	4	14.3	35	3	US-08-475-668A-70	Sequence 70, Appl
593	4	14.3	33	5	PCT-US95-04570-35	Sequence 35, Appl	666	4	14.3	35	3	US-08-475-668A-71	Sequence 71, Appl
594	4	14.3	33	5	PCT-US95-04589-34	Sequence 34, Appl	667	4	14.3	35	3	US-08-475-668A-72	Sequence 72, Appl
595	4	14.3	33	5	PCT-US95-10398-241	Sequence 241, App	668	4	14.3	35	3	US-08-485-551A-63	Sequence 63, Appl
596	4	14.3	33	5	PCT-US95-10398-242	Sequence 242, App	669	4	14.3	35	3	US-08-485-551A-64	Sequence 64, Appl
597	4	14.3	34	2	US-08-372-887-11	Sequence 11, Appl	670	4	14.3	35	3	US-08-485-551A-65	Sequence 65, Appl
598	4	14.3	34	2	US-08-690-011A-53	Sequence 53, Appl	671	4	14.3	35	3	US-08-485-551A-66	Sequence 66, Appl
599	4	14.3	34	2	US-08-690-011A-54	Sequence 54, Appl	672	4	14.3	35	3	US-08-485-551A-67	Sequence 67, Appl
600	4	14.3	34	2	US-08-690-011A-57	Sequence 57, Appl	673	4	14.3	35	3	US-08-485-551A-68	Sequence 68, Appl
601	4	14.3	35	1	US-08-127-351-15	Sequence 15, Appl	674	4	14.3	35	3	US-08-485-551A-69	Sequence 69, Appl
602	4	14.3	35	1	US-08-480-367B-15	Sequence 15, Appl	675	4	14.3	35	3	US-08-485-551A-70	Sequence 70, Appl
603	4	14.3	35	1	US-08-487-221A-15	Sequence 15, Appl	676	4	14.3	35	3	US-08-485-551A-71	Sequence 71, Appl
604	4	14.3	35	1	US-08-480-370-15	Sequence 15, Appl	677	4	14.3	35	3	US-08-485-551A-72	Sequence 72, Appl
605	4	14.3	35	2	US-08-488-161-22	Sequence 22, Appl	678	4	14.3	35	3	US-08-471-913A-63	Sequence 63, Appl
606	4	14.3	35	3	US-08-486-099-63	Sequence 63, Appl	679	4	14.3	35	3	US-08-471-913A-64	Sequence 64, Appl
607	4	14.3	35	3	US-08-486-099-64	Sequence 64, Appl	680	4	14.3	35	3	US-08-471-913A-65	Sequence 65, Appl
608	4	14.3	35	3	US-08-486-099-65	Sequence 65, Appl	681	4	14.3	35	3	US-08-471-913A-66	Sequence 66, Appl
609	4	14.3	35	3	US-08-486-099-66	Sequence 66, Appl	682	4	14.3	35	3	US-08-471-913A-67	Sequence 67, Appl
610	4	14.3	35	3	US-08-486-099-67	Sequence 67, Appl	683	4	14.3	35	3	US-08-471-913A-68	Sequence 68, Appl
611	4	14.3	35	3	US-08-486-099-68	Sequence 68, Appl	684	4	14.3	35	3	US-08-471-913A-69	Sequence 69, Appl

685	4	14.3	35	3	US-08-471-913A-70	Sequence 70, Appl	758	4	14.3	37	4	US-08-477-346-214	Sequence 214, App
686	4	14.3	35	3	US-08-471-913A-71	Sequence 71, Appl	759	4	14.3	37	4	US-08-473-089-169	Sequence 169, App
687	4	14.3	35	3	US-08-471-913A-72	Sequence 72, Appl	760	4	14.3	37	4	US-08-473-089-214	Sequence 214, App
688	4	14.3	35	4	US-08-927-219-43	Sequence 43, Appl	761	4	14.3	37	5	PCT-US95-09338-97	Sequence 97, Appl
689	4	14.3	35	4	US-08-485-264A-63	Sequence 63, Appl	762	4	14.3	37	5	PCT-US95-09339-97	Sequence 97, Appl
690	4	14.3	35	4	US-08-485-264A-64	Sequence 64, Appl	763	4	14.3	38	4	US-08-349-403-4	Sequence 4, Appl
691	4	14.3	35	4	US-08-485-264A-65	Sequence 65, Appl	764	4	14.3	40	1	US-07-868-353A-3	Sequence 3, Appl
692	4	14.3	35	4	US-08-485-264A-66	Sequence 66, Appl	765	4	14.3	40	2	US-08-407-804-3	Sequence 3, Appl
693	4	14.3	35	4	US-08-485-264A-67	Sequence 67, Appl	766	4	14.3	40	3	US-09-124-807-3	Sequence 3, Appl
694	4	14.3	35	4	US-08-485-264A-68	Sequence 68, Appl	767	4	14.3	41	6	5196523-6	Patent No. 5196523
695	4	14.3	35	4	US-08-485-264A-69	Sequence 69, Appl	768	4	14.3	42	1	US-08-520-599-1	Sequence 1, Appl
696	4	14.3	35	4	US-08-485-264A-70	Sequence 70, Appl	769	4	14.3	42	4	US-09-187-789-33	Sequence 33, Appl
697	4	14.3	35	4	US-08-485-264A-71	Sequence 71, Appl	770	4	14.3	43	2	US-08-488-161-40	Sequence 40, Appl
698	4	14.3	35	4	US-08-485-264A-72	Sequence 72, Appl	771	4	14.3	43	2	US-08-779-113-16	Sequence 16, Appl
699	4	14.3	35	4	US-09-082-279B-385	Sequence 385, App	772	4	14.3	43	3	US-09-273-685-40	Sequence 40, Appl
700	4	14.3	35	4	US-09-082-279B-386	Sequence 386, App	773	4	14.3	43	4	US-08-981-739-104	Sequence 104, App
701	4	14.3	35	4	US-09-082-279B-387	Sequence 387, App	774	4	14.3	43	5	PCT-US95-11934-40	Sequence 40, Appl
702	4	14.3	35	4	US-09-082-279B-388	Sequence 388, App	775	4	14.3	44	1	US-08-686-878A-27	Sequence 27, Appl
703	4	14.3	35	4	US-09-082-279B-389	Sequence 389, App	776	4	14.3	44	1	US-08-262-037-135	Sequence 135, App
704	4	14.3	35	4	US-09-082-279B-390	Sequence 390, App	777	4	14.3	44	2	US-08-292-870-2	Sequence 2, Appl
705	4	14.3	35	4	US-09-082-279B-391	Sequence 391, App	778	4	14.3	44	4	US-09-082-279B-51	Sequence 51, Appl
706	4	14.3	35	4	US-09-082-279B-392	Sequence 392, App	779	4	14.3	44	4	US-09-146-950-8	Sequence 8, Appl
707	4	14.3	35	4	US-09-082-279B-393	Sequence 393, App	780	4	14.3	44	4	US-09-146-950-24	Sequence 24, Appl
708	4	14.3	35	4	US-09-082-279B-394	Sequence 394, App	781	4	14.3	44	4	US-09-175-928-27	Sequence 27, Appl
709	4	14.3	35	4	US-09-082-279B-395	Sequence 395, App	782	4	14.3	44	4	US-09-315-304B-51	Sequence 51, Appl
710	4	14.3	35	4	US-09-082-279B-396	Sequence 396, App	783	4	14.3	45	2	US-08-637-759B-236	Sequence 236, App
711	4	14.3	35	4	US-08-749-066A-4	Sequence 4, Appl	784	4	14.3	45	2	US-08-779-113-41	Sequence 41, Appl
712	4	14.3	35	4	US-08-474-349A-63	Sequence 63, Appl	785	4	14.3	45	3	US-08-871-355A-236	Sequence 236, App
713	4	14.3	35	4	US-08-474-349A-64	Sequence 64, Appl	786	4	14.3	45	4	US-08-905-223-318	Sequence 318, App
714	4	14.3	35	4	US-08-474-349A-65	Sequence 65, Appl	787	4	14.3	45	4	US-08-776-059-9	Sequence 9, Appl
715	4	14.3	35	4	US-08-474-349A-66	Sequence 66, Appl	788	4	14.3	45	4	US-08-776-059-9	Sequence 9, Appl
716	4	14.3	35	4	US-08-474-349A-67	Sequence 67, Appl	789	4	14.3	45	4	US-08-776-059-10	Sequence 10, Appl
717	4	14.3	35	4	US-08-474-349A-68	Sequence 68, Appl	790	4	14.3	45	4	US-08-776-059-11	Sequence 11, Appl
718	4	14.3	35	4	US-08-474-349A-69	Sequence 69, Appl	791	4	14.3	45	4	US-08-630-915A-71	Sequence 71, Appl
719	4	14.3	35	4	US-08-474-349A-70	Sequence 70, Appl	792	4	14.3	45	4	US-09-201-945-236	Sequence 236, App
720	4	14.3	35	4	US-08-474-349A-71	Sequence 71, Appl	793	4	14.3	47	3	US-08-486-099-120	Sequence 120, App
721	4	14.3	35	4	US-08-474-349A-72	Sequence 72, Appl	794	4	14.3	47	3	US-08-360-107A-92	Sequence 92, Appl
722	4	14.3	35	4	US-08-474-349A-358	Sequence 358, App	795	4	14.3	47	3	US-08-360-107A-130	Sequence 130, App
723	4	14.3	35	4	US-08-474-349A-359	Sequence 359, App	796	4	14.3	47	3	US-08-360-107A-138	Sequence 138, App
724	4	14.3	35	4	US-08-474-349A-360	Sequence 360, App	797	4	14.3	47	3	US-08-484-223B-120	Sequence 120, App
725	4	14.3	35	4	US-08-474-349A-361	Sequence 361, App	798	4	14.3	47	3	US-08-919-597-120	Sequence 120, App
726	4	14.3	35	4	US-08-474-349A-362	Sequence 362, App	799	4	14.3	47	3	US-08-475-668A-120	Sequence 120, App
727	4	14.3	35	4	US-08-474-349A-363	Sequence 363, App	800	4	14.3	47	3	US-08-485-551A-120	Sequence 120, App
728	4	14.3	35	4	US-08-474-349A-364	Sequence 364, App	801	4	14.3	47	3	US-08-471-913A-120	Sequence 120, App
729	4	14.3	35	4	US-08-474-349A-365	Sequence 365, App	802	4	14.3	47	4	US-09-391-799-6	Sequence 6, Appl
730	4	14.3	35	4	US-08-474-349A-366	Sequence 366, App	803	4	14.3	47	4	US-08-485-264A-120	Sequence 120, App
731	4	14.3	35	4	US-08-474-349A-367	Sequence 367, App	804	4	14.3	47	4	US-08-474-349A-120	Sequence 120, App
732	4	14.3	35	4	US-08-474-349A-368	Sequence 368, App	805	4	14.3	48	4	US-09-314-268-95	Sequence 95, Appl
733	4	14.3	35	4	US-09-315-304B-385	Sequence 385, App	806	4	14.3	49	2	US-08-779-113-40	Sequence 40, Appl
734	4	14.3	35	4	US-09-315-304B-386	Sequence 386, App	807	4	14.3	49	4	US-09-004-406C-21	Sequence 21, Appl
735	4	14.3	35	4	US-09-315-304B-387	Sequence 387, App	808	4	14.3	50	4	US-08-905-223-378	Sequence 378, App
736	4	14.3	35	4	US-09-315-304B-388	Sequence 388, App	809	4	14.3	51	2	US-08-765-179B-3	Sequence 3, Appl
737	4	14.3	35	4	US-09-315-304B-389	Sequence 389, App	810	4	14.3	51	4	US-09-150-460B-16	Sequence 16, Appl
738	4	14.3	35	4	US-09-315-304B-390	Sequence 390, App	811	4	14.3	52	2	US-08-726-306A-143	Sequence 143, App
739	4	14.3	35	4	US-09-315-304B-391	Sequence 391, App	812	4	14.3	52	4	US-08-734-607B-21	Sequence 21, Appl
740	4	14.3	35	4	US-09-315-304B-392	Sequence 392, App	813	4	14.3	52	4	US-09-070-504-14	Sequence 14, Appl
741	4	14.3	35	4	US-09-315-304B-393	Sequence 393, App	814	4	14.3	53	2	US-08-779-113-13	Sequence 13, Appl
742	4	14.3	35	4	US-09-315-304B-394	Sequence 394, App	815	4	14.3	53	2	US-08-779-113-24	Sequence 24, Appl
743	4	14.3	35	4	US-09-315-304B-395	Sequence 395, App	816	4	14.3	53	2	US-08-779-113-25	Sequence 25, Appl
744	4	14.3	35	4	US-09-315-304B-396	Sequence 396, App	817	4	14.3	53	2	US-08-779-113-26	Sequence 26, Appl
745	4	14.3	35	5	PCT-US93-02011-2	Sequence 2, Appl	818	4	14.3	53	2	US-08-779-113-27	Sequence 27, Appl
746	4	14.3	35	5	PCT-US95-11934-22	Sequence 22, Appl	819	4	14.3	53	2	US-08-779-113-47	Sequence 47, Appl
747	4	14.3	36	4	US-09-082-279B-397	Sequence 397, App	820	4	14.3	53	2	US-08-779-113-48	Sequence 48, Appl
748	4	14.3	36	4	US-09-082-279B-859	Sequence 859, App	821	4	14.3	53	2	US-08-779-113-49	Sequence 49, Appl
749	4	14.3	36	4	US-09-315-304B-397	Sequence 397, App	822	4	14.3	53	2	US-08-779-113-50	Sequence 50, Appl
750	4	14.3	36	4	US-09-315-304B-859	Sequence 859, App	823	4	14.3	53	2	US-08-779-113-51	Sequence 51, Appl
751	4	14.3	37	1	US-08-190-802A-169	Sequence 169, App	824	4	14.3	53	4	US-09-091-814-92	Sequence 92, Appl
752	4	14.3	37	1	US-08-190-802A-214	Sequence 214, App	825	4	14.3	54	1	US-08-346-849-13	Sequence 13, Appl
753	4	14.3	37	2	US-08-631-328-49	Sequence 49, Appl	826	4	14.3	54	2	US-08-293-284A-13	Sequence 13, Appl
754	4	14.3	37	2	US-08-505-486-97	Sequence 97, Appl	827	4	14.3	54	3	US-08-917-299-12	Sequence 12, Appl
755	4	14.3	37	3	US-08-801-028-97	Sequence 97, Appl	828	4	14.3	54	3	US-08-917-299-13	Sequence 13, Appl
756	4	14.3	37	3	US-08-340-154-57	Sequence 97, Appl	829	4	14.3	54	3	US-08-917-299-14	Sequence 14, Appl
757	4	14.3	37	4	US-08-477-346-169	Sequence 169, App	830	4	14.3	54	3	US-08-917-299-15	Sequence 15, Appl

831	4	14.3	54	3	US-08-917-299-16	Sequence 16, Appl	904	4	14.3	61	2	US-08-829-876-134	Sequence 134, App
832	4	14.3	54	3	US-08-917-299-17	Sequence 17, Appl	905	4	14.3	61	2	US-08-829-876-135	Sequence 135, App
833	4	14.3	54	3	US-08-917-299-18	Sequence 18, Appl	906	4	14.3	61	2	US-08-829-876-136	Sequence 136, App
834	4	14.3	54	3	US-08-917-299-19	Sequence 19, Appl	907	4	14.3	61	2	US-08-829-876-137	Sequence 137, App
835	4	14.3	54	3	US-08-917-299-20	Sequence 20, Appl	908	4	14.3	61	2	US-08-829-876-138	Sequence 138, App
836	4	14.3	54	4	US-09-422-662-12	Sequence 12, Appl	909	4	14.3	61	2	US-08-829-876-139	Sequence 139, App
837	4	14.3	54	4	US-09-422-662-13	Sequence 13, Appl	910	4	14.3	61	2	US-08-829-876-140	Sequence 140, App
838	4	14.3	54	4	US-09-422-662-14	Sequence 14, Appl	911	4	14.3	61	2	US-08-829-876-141	Sequence 141, App
839	4	14.3	54	4	US-09-422-662-15	Sequence 15, Appl	912	4	14.3	61	2	US-08-829-876-142	Sequence 142, App
840	4	14.3	54	4	US-09-422-662-16	Sequence 16, Appl	913	4	14.3	61	2	US-08-829-876-143	Sequence 143, App
841	4	14.3	54	4	US-09-422-662-17	Sequence 17, Appl	914	4	14.3	61	2	US-08-829-876-144	Sequence 144, App
842	4	14.3	54	4	US-09-422-662-18	Sequence 18, Appl	915	4	14.3	61	2	US-08-829-876-145	Sequence 145, App
843	4	14.3	54	4	US-09-422-662-19	Sequence 19, Appl	916	4	14.3	61	2	US-08-829-876-146	Sequence 146, App
844	4	14.3	54	4	US-09-422-662-20	Sequence 20, Appl	917	4	14.3	61	2	US-08-829-876-147	Sequence 147, App
845	4	14.3	54	4	US-08-936-165A-377	Sequence 377, App	918	4	14.3	61	2	US-08-829-876-148	Sequence 148, App
846	4	14.3	55	1	US-08-394-326-7	Sequence 7, Appli	919	4	14.3	61	2	US-08-829-876-149	Sequence 149, App
847	4	14.3	55	3	US-09-082-306-7	Sequence 7, Appli	920	4	14.3	61	2	US-08-829-876-150	Sequence 150, App
848	4	14.3	55	3	US-08-814-836-7	Sequence 7, Appli	921	4	14.3	61	2	US-08-829-876-151	Sequence 151, App
849	4	14.3	55	4	US-09-220-528-103	Sequence 103, App	922	4	14.3	61	2	US-08-829-876-153	Sequence 153, App
850	4	14.3	56	6	5223482-30	Patent No. 5223482	923	4	14.3	61	2	US-08-829-876-154	Sequence 154, App
851	4	14.3	57	2	US-08-422-333-11	Sequence 11, Appl	924	4	14.3	61	2	US-08-829-876-155	Sequence 155, App
852	4	14.3	57	2	US-08-107-676-32	Sequence 32, Appl	925	4	14.3	61	2	US-08-829-876-156	Sequence 156, App
853	4	14.3	57	2	US-08-829-876-1	Sequence 1, Appli	926	4	14.3	61	2	US-08-829-876-157	Sequence 157, App
854	4	14.3	57	2	US-08-829-876-3	Sequence 3, Appli	927	4	14.3	61	2	US-08-829-876-158	Sequence 158, App
855	4	14.3	57	2	US-08-829-876-114	Sequence 114, App	928	4	14.3	61	2	US-08-829-876-159	Sequence 159, App
856	4	14.3	57	2	US-08-829-876-116	Sequence 116, App	929	4	14.3	61	2	US-08-829-876-160	Sequence 160, App
857	4	14.3	57	2	US-08-829-876-118	Sequence 118, App	930	4	14.3	61	2	US-08-829-876-161	Sequence 161, App
858	4	14.3	57	2	US-08-829-876-128	Sequence 128, App	931	4	14.3	61	2	US-08-829-876-163	Sequence 163, App
859	4	14.3	57	2	US-08-829-876-152	Sequence 152, App	932	4	14.3	61	2	US-08-829-876-164	Sequence 164, App
860	4	14.3	57	2	US-08-829-876-162	Sequence 162, App	933	4	14.3	61	2	US-08-829-876-166	Sequence 166, App
861	4	14.3	57	2	US-08-829-876-165	Sequence 165, App	934	4	14.3	61	2	US-08-829-876-167	Sequence 167, App
862	4	14.3	57	6	5187153-19	Patent No. 5187153	935	4	14.3	61	2	US-08-829-876-168	Sequence 168, App
863	4	14.3	57	6	5220013-21	Patent No. 5220013	936	4	14.3	61	2	US-08-829-876-169	Sequence 169, App
864	4	14.3	57	6	5220013-27	Patent No. 5220013	937	4	14.3	61	2	US-08-829-876-170	Sequence 170, App
865	4	14.3	57	6	5223482-21	Patent No. 5223482	938	4	14.3	61	2	US-08-829-876-171	Sequence 171, App
866	4	14.3	58	2	US-08-284-391B-35	Sequence 35, Appl	939	4	14.3	61	2	US-08-829-876-172	Sequence 172, App
867	4	14.3	58	2	US-08-829-876-5	Sequence 5, Appli	940	4	14.3	61	2	US-08-829-876-173	Sequence 173, App
868	4	14.3	58	3	US-09-053-197A-13	Sequence 13, Appl	941	4	14.3	61	2	US-08-829-876-174	Sequence 174, App
869	4	14.3	58	4	US-09-218-950-35	Sequence 35, Appl	942	4	14.3	61	2	US-08-829-876-175	Sequence 175, App
870	4	14.3	58	4	US-09-085-761A-13	Sequence 13, Appl	943	4	14.3	61	2	US-08-829-876-176	Sequence 176, App
871	4	14.3	59	2	US-08-829-876-6	Sequence 6, Appli	944	4	14.3	61	2	US-08-829-876-177	Sequence 177, App
872	4	14.3	59	4	US-09-177-249-164	Sequence 164, App	945	4	14.3	61	2	US-08-829-876-178	Sequence 178, App
873	4	14.3	60	4	US-09-015-030-9	Sequence 9, Appli	946	4	14.3	61	2	US-08-829-876-179	Sequence 179, App
874	4	14.3	61	1	US-08-682-485A-26	Sequence 26, Appl	947	4	14.3	61	2	US-08-829-876-180	Sequence 180, App
875	4	14.3	61	2	US-08-465-380-61	Sequence 61, Appl	948	4	14.3	61	2	US-08-829-876-181	Sequence 181, App
876	4	14.3	61	2	US-08-486-397-61	Sequence 61, Appl	949	4	14.3	61	2	US-08-829-876-182	Sequence 182, App
877	4	14.3	61	2	US-08-486-399-61	Sequence 61, Appl	950	4	14.3	61	2	US-08-829-876-183	Sequence 183, App
878	4	14.3	61	2	US-08-461-965-61	Sequence 61, Appl	951	4	14.3	61	2	US-08-829-876-184	Sequence 184, App
879	4	14.3	61	2	US-08-634-641-61	Sequence 61, Appl	952	4	14.3	61	2	US-08-829-876-185	Sequence 185, App
880	4	14.3	61	2	US-08-933-314-26	Sequence 26, Appl	953	4	14.3	61	2	US-08-829-876-186	Sequence 186, App
881	4	14.3	61	2	US-08-829-876-4	Sequence 4, Appli	954	4	14.3	61	2	US-08-829-876-187	Sequence 187, App
882	4	14.3	61	2	US-08-829-876-77	Sequence 77, Appl	955	4	14.3	61	2	US-08-829-876-188	Sequence 188, App
883	4	14.3	61	2	US-08-829-876-110	Sequence 110, App	956	4	14.3	61	2	US-08-829-876-189	Sequence 189, App
884	4	14.3	61	2	US-08-829-876-111	Sequence 111, App	957	4	14.3	61	2	US-08-829-876-190	Sequence 190, App
885	4	14.3	61	2	US-08-829-876-112	Sequence 112, App	958	4	14.3	61	2	US-08-829-876-191	Sequence 191, App
886	4	14.3	61	2	US-08-829-876-113	Sequence 113, App	959	4	14.3	61	2	US-08-829-876-192	Sequence 192, App
887	4	14.3	61	2	US-08-829-876-115	Sequence 115, App	960	4	14.3	61	2	US-08-829-876-193	Sequence 193, App
888	4	14.3	61	2	US-08-829-876-117	Sequence 117, App	961	4	14.3	61	2	US-08-829-876-194	Sequence 194, App
889	4	14.3	61	2	US-08-829-876-118	Sequence 118, App	962	4	14.3	61	2	US-08-829-876-195	Sequence 195, App
890	4	14.3	61	2	US-08-829-876-119	Sequence 119, App	963	4	14.3	61	2	US-08-829-876-196	Sequence 196, App
891	4	14.3	61	2	US-08-829-876-120	Sequence 120, App	964	4	14.3	61	2	US-08-829-876-197	Sequence 197, App
892	4	14.3	61	2	US-08-829-876-121	Sequence 121, App	965	4	14.3	61	2	US-08-829-876-198	Sequence 198, App
893	4	14.3	61	2	US-08-829-876-122	Sequence 122, App	966	4	14.3	61	2	US-08-829-876-199	Sequence 199, App
894	4	14.3	61	2	US-08-829-876-123	Sequence 123, App	967	4	14.3	61	2	US-08-829-876-200	Sequence 200, App
895	4	14.3	61	2	US-08-829-876-124	Sequence 124, App	968	4	14.3	61	2	US-08-829-876-201	Sequence 201, App
896	4	14.3	61	2	US-08-829-876-125	Sequence 125, App	969	4	14.3	61	2	US-08-829-876-202	Sequence 202, App
897	4	14.3	61	2	US-08-829-876-126	Sequence 126, App	970	4	14.3	61	2	US-08-829-876-203	Sequence 203, App
898	4	14.3	61	2	US-08-829-876-127	Sequence 127, App	971	4	14.3	61	2	US-08-829-876-204	Sequence 204, App
899	4	14.3	61	2	US-08-829-876-129	Sequence 129, App	972	4	14.3	61	2	US-08-829-876-205	Sequence 205, App
900	4	14.3	61	2	US-08-829-876-130	Sequence 130, App	973	4	14.3	61	2	US-08-829-876-206	Sequence 206, App
901	4	14.3	61	2	US-08-829-876-131	Sequence 131, App	974	4	14.3	61	2	US-08-829-876-207	Sequence 207, App
902	4	14.3	61	2	US-08-829-876-132	Sequence 132, App	975	4	14.3	61	2	US-08-829-876-208	Sequence 208, App
903	4	14.3	61	2	US-08-829-876-133	Sequence 133, App	976	4	14.3	61	2	US-08-829-876-209	Sequence 209, App

977	4	14.3	61	2	US-08-829-876-210	Sequence 210, App	1050	4	14.3	71	3	US-08-338-579A-103	Sequence 103, App
978	4	14.3	61	2	US-08-829-876-211	Sequence 211, App	1051	4	14.3	72	3	US-08-338-579A-96	Sequence 96, Appl
979	4	14.3	61	2	US-08-829-876-212	Sequence 212, App	1052	4	14.3	72	3	US-08-905-223-372	Sequence 372, App
980	4	14.3	61	2	US-08-829-876-213	Sequence 213, App	1053	4	14.3	73	1	US-08-274-366-3	Sequence 3, Appli
981	4	14.3	61	2	US-08-829-876-214	Sequence 214, App	1054	4	14.3	73	2	US-08-572-951-38	Sequence 38, Appl
982	4	14.3	61	2	US-08-829-876-215	Sequence 215, App	1055	4	14.3	73	2	US-08-392-816-19	Sequence 19, Appl
983	4	14.3	61	2	US-08-829-876-216	Sequence 216, App	1056	4	14.3	73	2	US-08-972-008-5	Sequence 5, Appli
984	4	14.3	61	2	US-08-829-876-217	Sequence 217, App	1057	4	14.3	73	3	US-08-735-491-6	Sequence 6, Appli
985	4	14.3	61	2	US-08-829-876-218	Sequence 218, App	1058	4	14.3	73	5	PCT-US95-07828-3	Sequence 3, Appli
986	4	14.3	61	2	US-08-829-876-219	Sequence 219, App	1059	4	14.3	74	2	US-08-450-905B-18	Sequence 18, Appl
987	4	14.3	61	2	US-08-829-876-220	Sequence 220, App	1060	4	14.3	74	2	US-08-450-905B-35	Sequence 35, Appl
988	4	14.3	61	2	US-08-829-876-221	Sequence 221, App	1061	4	14.3	74	3	US-08-941-445A-36	Sequence 36, Appl
989	4	14.3	61	2	US-08-829-876-222	Sequence 222, App	1062	4	14.3	75	3	US-09-100-804-20	Sequence 20, Appl
990	4	14.3	61	2	US-08-829-876-223	Sequence 223, App	1063	4	14.3	75	3	US-08-829-876-75	Sequence 75, Appl
991	4	14.3	61	2	US-08-829-876-224	Sequence 224, App	1064	4	14.3	78	4	US-08-858-207A-422	Sequence 422, App
992	4	14.3	61	2	US-08-829-876-225	Sequence 225, App	1065	4	14.3	78	5	PCT-US96-10521-6	Sequence 6, Appli
993	4	14.3	61	2	US-08-829-876-226	Sequence 226, App	1066	4	14.3	78	6	5187153-25	Patent No. 5187153
994	4	14.3	61	2	US-08-829-876-227	Sequence 227, App	1067	4	14.3	78	6	5187153-27	Patent No. 5187153
995	4	14.3	61	2	US-08-829-876-228	Sequence 228, App	1068	4	14.3	78	6	5223482-27	Patent No. 5223482
996	4	14.3	61	3	US-09-249-471-61	Sequence 61, Appl	1069	4	14.3	78	6	5223482-29	Patent No. 5223482
997	4	14.3	61	3	US-09-249-472-61	Sequence 61, Appl	1070	4	14.3	79	4	US-08-905-223-493	Sequence 493, App
998	4	14.3	61	3	US-09-249-451-61	Sequence 61, Appl	1071	4	14.3	79	4	US-08-858-207A-482	Sequence 482, App
999	4	14.3	61	3	US-08-809-455-61	Sequence 61, Appl	1072	4	14.3	83	3	US-08-341-018-12	Sequence 12, Appl
1000	4	14.3	61	3	US-08-809-455-61	Sequence 61, Appl	1073	4	14.3	83	3	US-08-470-335-198	Sequence 198, App
1001	4	14.3	61	3	US-09-249-461-61	Sequence 61, Appl	1074	4	14.3	83	4	US-09-382-155-16	Sequence 16, Appl
1002	4	14.3	61	3	US-09-249-448-61	Sequence 61, Appl	1075	4	14.3	83	4	US-09-074-044A-16	Sequence 16, Appl
1003	4	14.3	61	6	5187153-15	Patent No. 5187153	1076	4	14.3	83	4	US-08-470-339-198	Sequence 198, App
1004	4	14.3	62	1	US-08-530-550-23	Sequence 23, Appl	1077	4	14.3	83	4	US-08-981-988A-3	Sequence 3, Appli
1005	4	14.3	62	1	US-08-162-102C-41	Sequence 41, Appl	1078	4	14.3	83	4	US-08-981-988A-6	Sequence 6, Appli
1006	4	14.3	62	1	US-08-451-472-5	Sequence 5, Appli	1079	4	14.3	83	4	US-08-981-988A-8	Sequence 8, Appli
1007	4	14.3	62	4	US-09-187-789-17	Sequence 17, Appl	1080	4	14.3	83	4	US-08-981-988A-9	Sequence 9, Appli
1008	4	14.3	62	4	US-09-187-789-24	Sequence 24, Appl	1081	4	14.3	85	2	US-08-592-646A-18	Sequence 18, Appl
1009	4	14.3	62	5	PCT-US95-13660-23	Sequence 23, Appl	1082	4	14.3	86	4	US-08-905-223-280	Sequence 280, App
1010	4	14.3	62	6	5223482-18	Patent No. 5223482	1083	4	14.3	87	1	US-08-497-312-21	Sequence 21, Appl
1011	4	14.3	63	1	US-08-415-751-44	Sequence 44, Appl	1084	4	14.3	87	4	US-08-905-223-380	Sequence 380, App
1012	4	14.3	64	1	US-08-530-550-24	Sequence 24, Appl	1085	4	14.3	88	2	US-08-690-011A-11	Sequence 11, Appl
1013	4	14.3	64	1	US-08-530-550-25	Sequence 25, Appl	1086	4	14.3	89	1	US-08-446-038B-23	Sequence 23, Appl
1014	4	14.3	64	2	US-08-292-870-1	Sequence 1, Appli	1087	4	14.3	89	1	US-08-446-010B-23	Sequence 23, Appl
1015	4	14.3	64	5	PCT-US95-13660-24	Sequence 24, Appl	1088	4	14.3	89	2	US-08-805-445-23	Sequence 23, Appl
1016	4	14.3	64	5	PCT-US95-13660-25	Sequence 25, Appl	1089	4	14.3	89	2	US-08-064-067D-23	Sequence 23, Appl
1017	4	14.3	65	1	US-08-330-163-15	Sequence 15, Appl	1090	4	14.3	89	2	US-09-066-208-23	Sequence 23, Appl
1018	4	14.3	65	1	US-08-482-111-15	Sequence 15, Appl	1091	4	14.3	89	2	US-08-690-011A-52	Sequence 52, Appl
1019	4	14.3	65	2	US-07-808-457-1	Sequence 1, Appli	1092	4	14.3	89	4	US-09-087-465-20	Sequence 20, Appl
1020	4	14.3	65	2	US-07-808-457-7	Sequence 7, Appli	1093	4	14.3	91	1	US-08-480-449-22	Sequence 22, Appl
1021	4	14.3	65	2	US-07-808-457-9	Sequence 9, Appli	1094	4	14.3	91	2	US-08-660-542-22	Sequence 22, Appl
1022	4	14.3	65	2	US-08-829-876-79	Sequence 79, Appl	1095	4	14.3	91	2	US-07-808-457-11	Sequence 11, Appl
1023	4	14.3	65	4	US-09-627-376-5	Sequence 5, Appli	1096	4	14.3	91	2	US-07-808-457-21	Sequence 21, Appl
1024	4	14.3	65	5	PCT-US92-10178-1	Sequence 1, Appli	1097	4	14.3	91	3	US-08-814-836-18	Sequence 18, Appl
1025	4	14.3	65	5	PCT-US92-10178-7	Sequence 7, Appli	1098	4	14.3	91	4	US-08-905-223-421	Sequence 421, App
1026	4	14.3	65	5	PCT-US92-10178-9	Sequence 9, Appli	1099	4	14.3	91	4	US-08-479-603-22	Sequence 22, Appl
1027	4	14.3	65	6	5320958-17	Patent No. 5320958	1100	4	14.3	91	4	US-09-227-357-162	Sequence 162, App
1028	4	14.3	66	2	US-08-162-081B-44	Sequence 44, Appl	1101	4	14.3	91	5	PCT-US92-10178-11	Sequence 11, Appl
1029	4	14.3	66	2	US-08-780-872-44	Sequence 44, Appl	1102	4	14.3	91	5	PCT-US92-10178-21	Sequence 21, Appl
1030	4	14.3	66	4	US-09-085-957-44	Sequence 44, Appl	1103	4	14.3	92	1	US-07-792-988-2	Sequence 2, Appli
1031	4	14.3	66	4	US-09-461-997-312	Sequence 312, App	1104	4	14.3	92	1	US-08-347-492B-11	Sequence 11, Appl
1032	4	14.3	66	4	US-08-588-258B-10	Sequence 10, Appl	1105	4	14.3	92	1	US-08-375-346A-4	Sequence 4, Appli
1033	4	14.3	68	2	US-08-615-232A-10	Sequence 10, Appl	1106	4	14.3	92	2	US-08-428-713-6	Sequence 6, Appli
1034	4	14.3	68	3	US-08-470-323-10	Sequence 10, Appl	1107	4	14.3	92	2	US-08-421-144A-4	Sequence 4, Appli
1035	4	14.3	68	3	US-08-995-050-1	Sequence 1, Appli	1108	4	14.3	92	2	US-08-798-143-11	Sequence 11, Appl
1036	4	14.3	68	3	US-08-460-505-10	Sequence 10, Appl	1109	4	14.3	92	2	US-08-467-123B-4	Sequence 4, Appli
1037	4	14.3	68	5	PCT-US96-08295-10	Sequence 10, Appl	1110	4	14.3	92	3	US-08-904-179-6	Sequence 6, Appli
1038	4	14.3	69	1	US-08-280-443-23	Sequence 23, Appl	1111	4	14.3	92	3	US-08-808-720-10	Sequence 10, Appl
1039	4	14.3	69	1	US-08-457-459-23	Sequence 23, Appl	1112	4	14.3	92	4	US-08-928-383B-15	Sequence 15, Appl
1040	4	14.3	69	1	US-08-555-678-23	Sequence 23, Appl	1113	4	14.3	92	4	US-09-230-637-41	Sequence 41, Appl
1041	4	14.3	69	2	US-08-716-188-6	Sequence 6, Appli	1114	4	14.3	92	4	US-09-133-521-6	Sequence 6, Appli
1042	4	14.3	69	3	US-08-783-974-3	Sequence 3, Appli	1115	4	14.3	92	4	US-08-679-493A-157	Sequence 157, App
1043	4	14.3	69	3	US-07-982-759F-18	Sequence 18, Appl	1116	4	14.3	92	4	US-08-679-493A-158	Sequence 158, App
1044	4	14.3	69	3	US-07-982-759F-35	Sequence 35, Appl	1117	4	14.3	92	4	US-08-679-493A-159	Sequence 159, App
1045	4	14.3	69	4	US-08-981-988A-4	Sequence 4, Appli	1118	4	14.3	92	4	US-08-679-493A-160	Sequence 160, App
1046	4	14.3	69	5	PCT-US95-02275-23	Sequence 23, Appl	1119	4	14.3	92	4	US-08-936-165A-418	Sequence 418, App
1047	4	14.3	71	2	US-08-726-306A-37	Sequence 37, Appl	1120	4	14.3	93	1	US-09-230-371A-24	Sequence 24, Appl
1048	4	14.3	71	3	US-08-783-974-6	Sequence 6, Appli	1121	4	14.3	93	1	US-08-466-033-69	Sequence 69, Appl
1049	4	14.3	71	3	US-08-783-974-7	Sequence 7, Appli	1122	4	14.3	93	2	US-08-444-733-69	Sequence 69, Appl

1123	4	14.3	93	2	US-08-464-134-69	Sequence 69, Appl	1196	4	14.3	100	2	US-08-634-641-20	Sequence 20, Appl
1124	4	14.3	93	2	US-08-461-361-69	Sequence 69, Appl	1197	4	14.3	100	3	US-09-249-471-20	Sequence 20, Appl
1125	4	14.3	93	2	US-08-485-910-69	Sequence 69, Appl	1198	4	14.3	100	3	US-09-034-916-15	Sequence 15, Appl
1126	4	14.3	93	2	US-08-964-725-14	Sequence 14, Appl	1199	4	14.3	100	3	US-09-249-472-20	Sequence 20, Appl
1127	4	14.3	93	5	PCT-US95-06266-53	Sequence 53, Appl	1200	4	14.3	100	3	US-09-249-451-20	Sequence 20, Appl
1128	4	14.3	94	1	US-08-466-033-67	Sequence 67, Appl							
1129	4	14.3	94	2	US-08-444-733-67	Sequence 67, Appl							
1130	4	14.3	94	2	US-08-464-134-67	Sequence 67, Appl							
1131	4	14.3	94	2	US-08-461-361-67	Sequence 67, Appl							
1132	4	14.3	94	2	US-08-485-910-67	Sequence 67, Appl							
1133	4	14.3	94	2	US-07-728-215-35	Sequence 35, Appl							
1134	4	14.3	94	2	US-07-728-215-37	Sequence 37, Appl							
1135	4	14.3	94	3	US-08-297-395-10	Sequence 10, Appl							
1136	4	14.3	94	4	US-08-938-085A-35	Sequence 35, Appl							
1137	4	14.3	94	4	US-08-938-085A-37	Sequence 37, Appl							
1138	4	14.3	94	5	PCT-US95-06266-51	Sequence 51, Appl							
1139	4	14.3	95	1	US-07-956-700B-23	Sequence 23, Appl							
1140	4	14.3	95	1	US-08-476-537-23	Sequence 23, Appl							
1141	4	14.3	95	1	US-08-485-607-23	Sequence 23, Appl							
1142	4	14.3	95	2	US-08-475-879-23	Sequence 23, Appl							
1143	4	14.3	95	4	US-09-043-514-2	Sequence 2, Appl							
1144	4	14.3	95	4	US-09-227-357-598	Sequence 598, App							
1145	4	14.3	96	3	US-09-036-113-2	Sequence 2, Appl							
1146	4	14.3	96	4	US-08-465-343A-11	Sequence 11, Appl							
1147	4	14.3	96	4	US-08-945-983-9	Sequence 9, Appl							
1148	4	14.3	97	1	US-08-844-280-4	Sequence 4, Appl							
1149	4	14.3	97	2	US-08-467-822-40	Sequence 40, Appl							
1150	4	14.3	97	4	US-09-472-971-4	Sequence 4, Appl							
1151	4	14.3	97	4	US-09-006-726-4	Sequence 4, Appl							
1152	4	14.3	97	4	US-08-905-223-444	Sequence 444, App							
1153	4	14.3	97	4	US-08-432-697-40	Sequence 40, Appl							
1154	4	14.3	97	4	US-08-466-248-40	Sequence 40, Appl							
1155	4	14.3	98	1	US-07-795-859B-8	Sequence 8, Appl							
1156	4	14.3	98	1	US-08-211-202-118	Sequence 118, App							
1157	4	14.3	98	1	US-07-942-245-37	Sequence 37, Appl							
1158	4	14.3	98	1	US-08-457-616-8	Sequence 8, Appl							
1159	4	14.3	98	2	US-08-465-380-21	Sequence 21, Appl							
1160	4	14.3	98	2	US-08-480-478-49	Sequence 49, Appl							
1161	4	14.3	98	2	US-08-486-397-21	Sequence 21, Appl							
1162	4	14.3	98	2	US-08-486-399-21	Sequence 21, Appl							
1163	4	14.3	98	2	US-08-461-965-21	Sequence 21, Appl							
1164	4	14.3	98	2	US-08-428-197-48	Sequence 48, Appl							
1165	4	14.3	98	2	US-08-690-011A-19	Sequence 19, Appl							
1166	4	14.3	98	2	US-08-326-110A-49	Sequence 49, Appl							
1167	4	14.3	98	2	US-08-634-641-21	Sequence 21, Appl							
1168	4	14.3	98	2	US-08-665-202-31	Sequence 31, Appl							
1169	4	14.3	98	3	US-09-249-471-21	Sequence 21, Appl							
1170	4	14.3	98	3	US-09-249-472-21	Sequence 21, Appl							
1171	4	14.3	98	3	US-09-249-451-21	Sequence 21, Appl							
1172	4	14.3	98	3	US-08-809-455-21	Sequence 21, Appl							
1173	4	14.3	98	3	US-09-249-461-21	Sequence 21, Appl							
1174	4	14.3	98	3	US-09-249-448-21	Sequence 21, Appl							
1175	4	14.3	98	5	PCT-US93-10555-48	Sequence 48, Appl							
1176	4	14.3	99	1	US-08-085-122-4	Sequence 4, Appl							
1177	4	14.3	99	2	US-08-319-052-4	Sequence 4, Appl							
1178	4	14.3	99	2	US-08-710-749-9	Sequence 9, Appl							
1179	4	14.3	99	2	US-08-710-749-10	Sequence 10, Appl							
1180	4	14.3	99	2	US-08-710-749-11	Sequence 11, Appl							
1181	4	14.3	99	2	US-08-710-749-13	Sequence 13, Appl							
1182	4	14.3	99	2	US-08-710-749-15	Sequence 15, Appl							
1183	4	14.3	99	4	US-08-931-858E-170	Sequence 170, App							
1184	4	14.3	99	4	US-08-905-223-278	Sequence 278, App							
1185	4	14.3	99	4	US-08-905-223-492	Sequence 492, App							
1186	4	14.3	99	4	US-08-981-739-170	Sequence 170, App							
1187	4	14.3	99	4	US-08-442-108B-4	Sequence 4, Appl							
1188	4	14.3	100	1	US-08-440-377A-3	Sequence 3, Appl							
1189	4	14.3	100	2	US-08-687-852-3	Sequence 3, Appl							
1190	4	14.3	100	2	US-08-465-380-20	Sequence 20, Appl							
1191	4	14.3	100	2	US-08-480-478-48	Sequence 48, Appl							
1192	4	14.3	100	2	US-08-486-397-20	Sequence 20, Appl							
1193	4	14.3	100	2	US-08-486-399-20	Sequence 20, Appl							
1194	4	14.3	100	2	US-08-461-965-20	Sequence 20, Appl							
1195	4	14.3	100	2	US-08-326-110A-48	Sequence 48, Appl							

ALIGNMENTS

RESULT 1
US-09-088-425-2
; Sequence 2, Application US/09088425
; Patent No. 6171843
; GENERAL INFORMATION:
; APPLICANT: BANDMAN, OLGA
; APPLICANT: LAL, PREETI
; APPLICANT: CORLEY, NEIL C.
; APPLICANT: PATTERSON, CHANDRA
; APPLICANT: BAUGHN, MARIAH R.
; TITLE OF INVENTION: HUMAN ISOMERASE-LIKE PROTEINS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/088,425
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Certione, Michael C
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0529 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 443 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: TMLR3DT01
CLONE: 289973
US-09-088-425-2

Query Match 21.4%; Score 6; DB 4; Length 443;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 QGLRAE 25
Db 286 QGLRAE 291
RESULT 2
US-08-980-523-9
; Sequence 9, Application US/08980523

Patent No. 6310181
GENERAL INFORMATION:
APPLICANT: Kouhara, Haruhiko
APPLICANT: Spivak-Kroizman, Taly
APPLICANT: Lax, Irit
APPLICANT: Schllessinger, Joseph
TITLE OF INVENTION: ADAPTOR PROTEIN FR52 AND
RELATED PRODUCTS AND METHODS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,523
FILING DATE: December 1, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/21851
FILING DATE: December 1, 1997
APPLICATION NUMBER: 60/032,093
FILING DATE: December 3, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 230/045
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 508 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-980-523-9

Query Match 21.4%; Score 6; DB 4; Length 508;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ENLPSL 16
Db 350 ENLPSL 355

RESULT 3
US-07-683-957B-3
Sequence 3, Application US/07683957B
Patent No. 5310880
GENERAL INFORMATION:
APPLICANT: Donahoe, Patricia K.
APPLICANT: Ragin, Richard C.
APPLICANT: MacLaughlin, David T.
TITLE OF INVENTION: Purification of M llerian Inhibiting
TITLE OF INVENTION: Substance
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, N.W.

CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/683,957B
FILING DATE: 19910412
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Goldstein, Jorge A.
REGISTRATION NUMBER: 29,021
REFERENCE/DOCKET NUMBER: 0609.3060000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 466-0800
TELEFAX: (202) 833-8716
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 553 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-683-957B-3

Query Match 21.4%; Score 6; DB 1; Length 553;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
Db 428 QGLRAE 433

RESULT 4
US-07-683-957B-2
Sequence 2, Application US/07683957B
Patent No. 5310880
GENERAL INFORMATION:
APPLICANT: Donahoe, Patricia K.
APPLICANT: Ragin, Richard C.
APPLICANT: MacLaughlin, David T.
TITLE OF INVENTION: Purification of M llerian Inhibiting
TITLE OF INVENTION: Substance
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/683,957B
FILING DATE: 19910412
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Goldstein, Jorge A.
REGISTRATION NUMBER: 29,021
REFERENCE/DOCKET NUMBER: 0609.3060000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 466-0800
TELEFAX: (202) 833-8716
INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 575 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-683-957B-2

Query Match 21.4%; Score 6; DB 1; Length 575;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 OGLRAE 25
Db 447 OGLRAE 452

RESULT 5
US-08-947-965-74
; Sequence 74, Application US/08947965A
; Patent No. 6004790
; GENERAL INFORMATION:
; APPLICANT: Dijkhuizen, Lubbert
; APPLICANT: Dijkstra, Bauke
; APPLICANT: Andersen, Carsten
; APPLICANT: Osten, Claus von der
; TITLE OF INVENTION: Cyclomaltodextrin Glucanotransferase
; TITLE OF INVENTION: Variants
; FILE REFERENCE: 4285.204-US
; CURRENT APPLICATION NUMBER: US/08/947,965A
; CURRENT FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 0477/95
; EARLIER FILING DATE: 1995-04-21
; EARLIER APPLICATION NUMBER: 1173/95
; EARLIER FILING DATE: 1995-10-17
; EARLIER APPLICATION NUMBER: 1281/95
; EARLIER FILING DATE: 1995-11-16
; EARLIER APPLICATION NUMBER: PCT/DK96/00179
; EARLIER FILING DATE: 1996-04-22
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 74
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-08-947-965-74

Query Match 21.4%; Score 6; DB 3; Length 685;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
Db 537 TGSALT 542

RESULT 6
US-09-368-590-2
; Sequence 2, Application US/09368590
; Patent No. 6187563
; GENERAL INFORMATION:
; APPLICANT: Solimena, Michele
; TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
; TITLE OF INVENTION: AUTOANTIGENS OF AUTOIMMUNE DISEASES
; FILE REFERENCE: 101918-200 (OCR-941)
; CURRENT APPLICATION NUMBER: US/09/368,590
; CURRENT FILING DATE: 1999-08-04
; EARLIER APPLICATION NUMBER: 60/095,657
; EARLIER FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2

; LENGTH: 2293
; TYPE: PRT
; ORGANISM: Human
US-09-368-590-2

Query Match 21.4%; Score 6; DB 4; Length 2293;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GLRAEE 26
Db 1077 GLRAEE 1082

RESULT 7
US-07-683-957B-4
; Sequence 4, Application US/07683957B
; Patent No. 5310880
; GENERAL INFORMATION:
; APPLICANT: Donahoe, Patricia K.
; APPLICANT: Ragin, Richard C.
; APPLICANT: MacLaughlin, David T.
; TITLE OF INVENTION: Purification of M llerian Inhibiting
; TITLE OF INVENTION: Substance
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/683,957B
; FILING DATE: 19910412
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldstein, Jorge A.
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 0609.3060000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 466-0800
; TELEFAX: (202) 833-8716
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-683-957B-4

Query Match 17.9%; Score 5; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 LRAEE 26
Db 1 LRAEE 5

RESULT 8
US-08-447-411-16
; Sequence 16, Application US/08447411
; Patent No. 5773243
; GENERAL INFORMATION:
; APPLICANT: FRITZINGER, DAVID C.

APPLICANT: BREDEHORST, REINHARD
APPLICANT: VOGEL, CARL-WILHELM
TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,411
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/043,747
FILING DATE: 07-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5773243man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-101-0
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-447-411-16

Query Match 17.9%; Score 5; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 AEENI 28
|||||
Db 6 AEENI 10

RESULT 9
US-08-447-411-48
Sequence 48, Application US/08447411
Patent No. 5773243
GENERAL INFORMATION:
APPLICANT: FRITZINGER, DAVID C.
APPLICANT: BREDEHORST, REINHARD
APPLICANT: VOGEL, CARL-WILHELM
TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,411
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/043,747
FILING DATE: 07-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5773243man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-101-0
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-447-411-48
Query Match 17.9%; Score 5; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 24 AEENI 28
|||||
Db 6 AEENI 10
RESULT 10
US-08-662-227-5
Sequence 5, Application US/08662227
Patent No. 5922320
GENERAL INFORMATION:
APPLICANT: VOGEL, CARL-WILHELM
APPLICANT: BREDEHORST, REINHORST
APPLICANT: KOCK, MICHAEL
APPLICANT: FRITZINGER, DAVID
TITLE OF INVENTION: RECOMBINANT PROCVF
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/662,227
FILING DATE: 14-JUN-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-0107-0X
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-662-227-5

Query Match 17.9%; Score 5; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 AEENI 28
| | | | |
Db 6 AEENI 10

RESULT 11
US-09-017-947-5
; Sequence 5, Application US/09017947
; Patent No. 6303754
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; APPLICANT: BREDEHORST, REINHARDT
; APPLICANT: KOCK, MICHAEL
; APPLICANT: FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/017,947
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/562,227
; FILING DATE: 14-JUN-1996

ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-017-947-5

Query Match 17.9%; Score 5; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 AEENI 28
| | | | |
Db 6 AEENI 10

RESULT 12
PCT-US95-04121-3
; Sequence 3, Application PC/TUS9504121
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Haptenated Peptides and Uses Thereof
; NUMBER OF SEQUENCES: 62
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04121
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/222,206
; FILING DATE: April 1, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Vanstone, Darlene A.
; REGISTRATION NUMBER: 35,279
; REFERENCE/DOCKET NUMBER: 079.2PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6010
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
PCT-US95-04121-3

Query Match 17.9%; Score 5; DB 5; Length 13;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 AEENI 28
| | | | |
Db 5 AEENI 9

RESULT 13
US-08-447-411-20
; Sequence 20, Application US/08447411
; Patent No. 5773243
; GENERAL INFORMATION:
; APPLICANT: FRITZINGER, DAVID C.
; APPLICANT: BREDEHORST, REINHARD
; APPLICANT: VOGEL, CARL-WILHELM
; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,411

;; FILING DATE: 435
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/043,747
;; FILING DATE: 07-APR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Oblon, No. 5773243man F.
;; REGISTRATION NUMBER: 24,618
;; REFERENCE/DOCKET NUMBER: 1126-101-0
;; TELEPHONE: (703) 413-3000
;; TELEFAX: (703) 413-2220
;; TELEX: 248855 OPAT UR
;; INFORMATION FOR SEQ ID NO: 20:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 21 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; ORIGINAL SOURCE:
;; ORGANISM: Homo sapiens
US-08-447-411-20

Query Match 17.9%; Score 5; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 AEENI 28

Db 9 AEENI 13

RESULT 14

US-08-447-411-60
; Sequence 60, Application US/08447411
; Patent No. 5773243
; GENERAL INFORMATION:
; APPLICANT: FRITZINGER, DAVID C.
; APPLICANT: BREDEHORST, REINHARD
; APPLICANT: VOGEL, CARL-WILHELM
; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,411
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/043,747
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5773243man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-101-0
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:

;; LENGTH: 21 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; ORIGINAL SOURCE:
;; ORGANISM: Homo sapiens
US-08-447-411-60

Query Match 17.9%; Score 5; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 AEENI 28

Db 9 AEENI 13

RESULT 15

US-08-662-227-17
; Sequence 17, Application US/08662227
; Patent No. 5922320
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; APPLICANT: BREDEHORST, REINHORST
; APPLICANT: KOCK, MICHAEL
; APPLICANT: FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/662,227
; FILING DATE: 14-JUN-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-662-227-17

QY 24 AEENI 28

Db 9 AEENI 13

RESULT 16

Query Match 17.9%; Score 5; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-017-947-17
; Sequence 17, Application US/09017947
; Patent No. 6303754
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; APPLICANT: BREDEHORST, REINHORST
; APPLICANT: KOCK, MICHAEL
; APPLICANT: FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/017.947
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/662,227
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-017-947-17

Query Match 17.9%; Score 5; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 AEENI 28
Db 9 AEENI 13

RESULT 17
US-08-447-411-73
; Sequence 73, Application US/08447411
; Patent No. 5773243
; GENERAL INFORMATION:
; APPLICANT: FRITZINGER, DAVID C.
; APPLICANT: BREDEHORST, REINHARD
; APPLICANT: VOGEL, CARL-WILHELM
; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.

ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,411
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/043,747
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, No. 5773243man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-101-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-08-447-411-73

Query Match 17.9%; Score 5; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 AEENI 28
Db 9 AEENI 13

RESULT 18
US-08-662-227-31
; Sequence 31, Application US/08662227
; Patent No. 5922320
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; APPLICANT: BREDEHORST, REINHORST
; APPLICANT: KOCK, MICHAEL
; APPLICANT: FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/662,227
; FILING DATE: 14-JUN-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-0X

TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-662-227-31

Query Match 17.9%; Score 5; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 AENI 28
Db 9 AENI 13

RESULT 19
US-09-017-947-31
Sequence 31, Application US/09017947
Patent No. 6303754
GENERAL INFORMATION:
APPLICANT: VOGEL, CARL-WILHELM
APPLICANT: BREDEHORST, REINHORST
APPLICANT: KOCK, MICHAEL
APPLICANT: FRITZINGER, DAVID
TITLE OF INVENTION: RECOMBINANT PROCVF
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/017,947
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/662,227
FILING DATE: 14-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-0107-0X
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-017-947-31

Query Match 17.9%; Score 5; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 AENI 28
Db 9 AENI 13
RESULT 20
US-08-145-708A-4
Sequence 4, Application US/08145708A
Patent No. 5519114
GENERAL INFORMATION:
APPLICANT: Johnson, Howard M.
APPLICANT: Torres, Barbara A.
APPLICANT: Yamamoto, Janet K.
TITLE OF INVENTION: Retroviral Superantigens, Superantigen
TITLE OF INVENTION: Peptides, and Methods of Use
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/145,708A
FILING DATE: 29-OCT-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF135
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-145-708A-4

Query Match 17.9%; Score 5; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLR 24
Db 15 QGLR 19

RESULT 21
US-08-331-454-4
Sequence 4, Application US/08331454
Patent No. 5968514
GENERAL INFORMATION:
APPLICANT: Johnson, Howard M.
APPLICANT: Torres, Barbara A.
APPLICANT: Yamamoto, Janet K.
TITLE OF INVENTION: Retroviral Superantigens, Superantigen
TITLE OF INVENTION: Peptides, and Methods of Use
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1

; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/331,454
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/145,708
; FILING DATE: 29-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF135
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-331-454-4

Query Match 17.9%; Score 5; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 OGLRA 24
|
|
|
|
Db 15 OGLRA 19

RESULT 22
US-09-187-789-43
; Sequence 43, Application US/09187789
; Patent No. 6340740
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; APPLICANT: Fernandez-Alnemri, Teresa
; TITLE OF INVENTION: CASPASE-14, AN APOPTOTIC PROTEASE, NUCLEIC ACID ENCODING
; FILE REFERENCE: 480140.434C1
; CURRENT APPLICATION NUMBER: US/09/187,789
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-187-789-43

Query Match 17.9%; Score 5; DB 4; Length 47;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 LRAEE 26
|
|
|
|
Db 33 LRAEE 37

RESULT 23
US-08-936-165A-369
; Sequence 369, Application US/08936165A
; Patent No. 6348582
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
; TITLE OF INVENTION: Polypeptides and Their Uses
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,165A
; FILING DATE: 24-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,032
; FILING DATE: 24-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R.
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 369:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 88 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; US-08-936-165A-369

Query Match 17.9%; Score 5; DB 4; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GSAIT 6
|
|
|
|
Db 56 GSAIT 60

RESULT 24
US-08-417-460-10
; Sequence 10, Application US/08417460
; Patent No. 5723754
; GENERAL INFORMATION:
; APPLICANT: Scott, Roderick J.
; APPLICANT: Draper, John
; APPLICANT: Wyatt, Paul
; TITLE OF INVENTION: TAPETUM SPECIFIC PROMOTERS FROM BRASSICACEAE SPP

NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hale and Dorr
STREET: 1455 Pennsylvania Ave. NW
CITY: Washington
STATE: D.C.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/417,460
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/078,228
FILING DATE: 23 August 1993
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-942-8400
TELEFAX: 202-942-8484
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 96 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
OTHER INFORMATION: product= "At A9 amino acid"

US-08-417-460-10

Query Match 17.9%; Score 5; DB 1; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 NLP5L 16
|||||
Db 86 NLP5L 90

RESULT 25
US-08-477-347-9
Sequence 9, Application US/08477347
Patent No. 623246
GENERAL INFORMATION:
APPLICANT: WALLACH, David
APPLICANT: BIGDA, Jacek
APPLICANT: BELETSKY, Igor
APPLICANT: MEYI, Igor
TITLE OF INVENTION: TNF LIGANDS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEWMARK
STREET: 419 Seventh Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,347
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/115,685
FILING DATE:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: IL 106271
FILING DATE: 08-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, G. Kevin
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: WALLACH-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 102 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-477-347-9

Query Match 17.9%; Score 5; DB 4; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LPSLC 17
|||||
Db 84 LPSLC 88

RESULT 26
US-08-850-910A-30
Sequence 30, Application US/08850910A
Patent No. 5948761
GENERAL INFORMATION:
APPLICANT: SEILHAMER, J.J.
APPLICANT: LEWICKI, J.
APPLICANT: SCARBOROUGH, R.M.
TITLE OF INVENTION: RECOMBINANT TECHNIQUES FOR
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER, LLP
STREET: 2000 Pennsylvania Avenue, NW, Suite 5500
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/850,910A
FILING DATE: 05-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/477,226
FILING DATE: 08-FEB-1990
APPLICATION NUMBER: 07/299,880
FILING DATE: 19-JAN-1989
APPLICATION NUMBER: 07/206,470
FILING DATE: 14-JUN-1988
APPLICATION NUMBER: 07/200,383
FILING DATE: 31-MAY-1988
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2190002025212
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-887-1500
TELEFAX: 202-822-0168
TELEX:
INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:
LENGTH: 106 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-850-910A-30

Query Match 17.9%; Score 5; DB 2; Length 106;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 QGLRA 24
|
|
|
|
Db 25 QGLRA 29

RESULT 27
US-09-460-384-34
; Sequence 34, Application US/09460384
; Patent No. 6337316
; GENERAL INFORMATION:
; APPLICANT: EL TAYAR, Nabil
; BLECHNER, Steven
; JAMESON, Brad
; TEPPER, Mark
; TITLE OF INVENTION: CD28/CTLA-4 INHIBITING PEPTIDOMIMETICS,
; PHARMACEUTICAL COMPOSITIONS THEREOF, AND METHOD OF USING
; SAME
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 624 Ninth Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/460,384
; FILING DATE: 13-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/12312
; FILING DATE: 11-JUN-1998
; APPLICATION NUMBER: US 60/049,470
; FILING DATE: 12-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: YUN, Allen C
; REGISTRATION NUMBER: 37,971
; REFERENCE/DOCKET NUMBER: EL TAYAR-1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-09-460-384-34

Query Match 17.9%; Score 5; DB 4; Length 109;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 20 QGLRA 24
|
|
|
|
Db 80 QGLRA 84

RESULT 28
US-08-630-172-4
; Sequence 4, Application US/08630172
; Patent No. 6060054
; GENERAL INFORMATION:
; APPLICANT: Staerz, Uwe
; TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T
; TITLE OF INVENTION: LYMPHOCYTE VETO
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, 35th Floor
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,172
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2879-36
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 124 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-630-172-4

Query Match 17.9%; Score 5; DB 3; Length 124;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 QGLRA 24
|
|
|
|
Db 80 QGLRA 84

RESULT 29
US-09-375-419-4
; Sequence 4, Application US/09375419
; Patent No. 6264950
; GENERAL INFORMATION:
; APPLICANT: Staerz, Uwe
; TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T
; TITLE OF INVENTION: LYMPHOCYTE VETO
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, 35th Floor
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.
; ZIP: 80203

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/375,419
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/630,172
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2879-36
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 124 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-375-419-4

Query Match 17.9% Score 5; DB 4; Length 124;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRA 24
| | | | |
Db 80 QGLRA 84

RESULT 30
US-08-170-360-3
; Sequence 3, Application US/08170360
; Patent No. 5656602
; GENERAL INFORMATION:
; APPLICANT: Tseng, Albert P. S.
; APPLICANT: Inglis, Adam
; TITLE OF INVENTION: PLA2 INHIBITORY COMPOUNDS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg Ernst & Kurz
; STREET: Suite 701-E, 555 Thirteenth St., N.W
; CITY: Washington
; STATE: D. C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/170,360
; FILING DATE: 03-MAR-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PCT/AU92/00333
; FILING DATE: 06-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PK 7058
; FILING DATE: 04-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Ernst, Barbara G.
; REGISTRATION NUMBER: 30,377
; REFERENCE/DOCKET NUMBER: 1871-104A
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)783-6040
; TELEFAX: (202)783-6031
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; US-08-170-360-3

Query Match 17.9% Score 5; DB 1; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GSAIT 6
| | | | |
Db 79 GSAIT 83

RESULT 31
US-08-888-497-38
; Sequence 38, Application US/08888497
; Patent No. 5972677
; GENERAL INFORMATION:
; APPLICANT: Tischfield, Jay A.
; APPLICANT: Sellhamer, Jeffrey J.
; TITLE OF INVENTION: Mammalian Phospholipase A2 Nucleotide
; TITLE OF INVENTION: Sequences and Low Molecular Weight Amino Acid Sequences
; TITLE OF INVENTION: Encoded Thereby, Antisense Sequences and Nucleotide
; TITLE OF INVENTION: Sequences Having Internal Ribosome Binding Sites
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &
; STREET: 200 East Broward Boulevard
; CITY: Fort Lauderdale
; STATE: FL
; COUNTRY: USA
; ZIP: 33301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/888,497
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/651,405
; FILING DATE:
; APPLICATION NUMBER: US 08/097,354
; FILING DATE: 26-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Manso, Peter J.
; REGISTRATION NUMBER: 32,264
; REFERENCE/DOCKET NUMBER: IN21044-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 305-527-2498
; TELEFAX: 305-764-4996
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-888-497-38

Query Match 17.9%; Score 5; DB 2; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GSAIT 6
Db 79 GSAIT 83

RESULT 32

US-09-362-230-38
; Sequence 38, Application US/09362230
; Patent No. 6352849

; GENERAL INFORMATION:

; APPLICANT: Tischfield, Jay A.
; TITLE OF INVENTION: Mammalian Phospholipase A2 Nucleotide
; TITLE OF INVENTION: Sequences and Low Molecular Weight Amino Acid Sequences
; TITLE OF INVENTION: Encoded Thereby, Antisense Sequences and Nucleotide
; TITLE OF INVENTION: Sequences Having Internal Ribosome Binding Sites
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &
; ADDRESS: Russell PA
; STREET: 200 East Broward Boulevard
; CITY: Fort Lauderdale
; STATE: FL
; COUNTRY: USA
; ZIP: 33301

COMPUTER READABLE FORM:

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/362,230
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/888,497
; FILING DATE:
; APPLICATION NUMBER: US 08/097,354
; FILING DATE: 26-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Manso, Peter J.
; REGISTRATION NUMBER: 32,264
; REFERENCE/DOCKET NUMBER: IN21044-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 305-527-2498
; TELEFAX: 305-764-4996
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US94-07926-38

Query Match 17.9%; Score 5; DB 4; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GSAIT 6
Db 79 GSAIT 83

RESULT 33

PCT-US94-07926-38

; Sequence 38, Application PC/TUS9407926

; GENERAL INFORMATION:

; APPLICANT: Tischfield, Jay A.
; TITLE OF INVENTION: Mammalian Phospholipase A2 Nucleotide
; TITLE OF INVENTION: Sequences and Low Molecular Weight Amino Acid Sequences
; TITLE OF INVENTION: Encoded Thereby, Antisense Sequences and Nucleotide
; TITLE OF INVENTION: Sequences Having Internal Ribosome Binding Sites
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &
; ADDRESS: Russell PA
; STREET: 200 East Broward Boulevard
; CITY: Fort Lauderdale
; STATE: FL
; COUNTRY: USA
; ZIP: 33301

COMPUTER READABLE FORM:

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07926
; FILING DATE: 15-JUL-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/097,354
; FILING DATE: 26-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Manso, Peter J.
; REGISTRATION NUMBER: 32,264
; REFERENCE/DOCKET NUMBER: IN21044-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 305-527-2498
; TELEFAX: 305-764-4996
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US94-07926-38

Query Match 17.9%; Score 5; DB 5; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GSAIT 6
Db 79 GSAIT 83

RESULT 34

US-08-804-180C-2
; Sequence 2, Application US/08804180C
; Patent No. 6107056

; GENERAL INFORMATION:

; APPLICANT: Martin K. Oaks
; TITLE OF INVENTION: sCTLA-4 and Its Soluble Products
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thomas M. Wozny
; STREET: 100 East Wisconsin Avenue
; CITY: Milwaukee
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53202

COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.50 inch Disk
; COMPUTER: IBM

; OPERATING SYSTEM: DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,180C
; FILING DATE: February 20, 1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Thomas M. Wozny
; REGISTRATION NUMBER: 28,922
; REFERENCE/DOCKET NUMBER: 3284-00003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 271-7590
; TELEFAX: (414) 271-5770
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 137
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; DESCRIPTION: peptide
; HYPOTHETICAL: no
; FRAGMENT TYPE: Mature Polypeptide
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapien
; DEVELOPMENTAL STAGE: Adult
; TISSUE TYPE: Lymphnode
; FEATURE:
; NAME/KEY: Human sCTLA-4
; IDENTIFICATION METHOD: Found by experiment
; OTHER INFORMATION: Asn 76 and Asn 108 represent N-linked glycosylation; B7 binding site
US-08-804-180C-2

Query Match 17.9%; Score 5; DB 3; Length 137;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 OQLRA 24
|||||
DB 80 OQLRA 84

RESULT 35
US-08-888-497-36
; Sequence 36, Application US/08888497
; Patent No. 5972677
; GENERAL INFORMATION:
; APPLICANT: Tischfield, Jay A.
; APPLICANT: Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Mammalian Phospholipase A2 Nucleotide
; TITLE OF INVENTION: Sequences and Low Molecular Weight Amino Acid Sequences
; TITLE OF INVENTION: Encoded Thereby, Antisense Sequences and Nucleotide
; TITLE OF INVENTION: Sequences Having Internal Ribosome Binding Sites
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &
; ADDRESSEE: Russell PA
; STREET: 200 East Broward Boulevard
; CITY: Fort Lauderdale
; STATE: FL
; COUNTRY: USA
; ZIP: 33301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/888,497
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/651,405

; FILING DATE:
; APPLICATION NUMBER: US 08/097,354
; FILING DATE: 26-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Manso, Peter J.
; REGISTRATION NUMBER: 32,264
; REFERENCE/DOCKET NUMBER: IN21044-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 305-527-2498
; TELEFAX: 305-764-4996
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 148 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-888-497-36

Query Match 17.9%; Score 5; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GSAIT 6
|||||
DB 101 GSAIT 105

RESULT 36
US-09-362-230-36
; Sequence 36, Application US/09362230
; Patent No. 6352849
; GENERAL INFORMATION:
; APPLICANT: Tischfield, Jay A.
; APPLICANT: Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Mammalian Phospholipase A2 Nucleotide
; TITLE OF INVENTION: Sequences and Low Molecular Weight Amino Acid Sequences
; TITLE OF INVENTION: Encoded Thereby, Antisense Sequences and Nucleotide
; TITLE OF INVENTION: Sequences Having Internal Ribosome Binding Sites
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &
; ADDRESSEE: Russell PA
; STREET: 200 East Broward Boulevard
; CITY: Fort Lauderdale
; STATE: FL
; COUNTRY: USA
; ZIP: 33301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/362,230
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/888,497
; FILING DATE:
; APPLICATION NUMBER: US 08/097,354
; FILING DATE: 26-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Manso, Peter J.
; REGISTRATION NUMBER: 32,264
; REFERENCE/DOCKET NUMBER: IN21044-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 305-527-2498
; TELEFAX: 305-764-4996
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 148 amino acids

;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-362-230-36

Query Match 17.9%; Score 5; DB 4; Length 148;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GSAIT 6
|||||
Db 101 GSAIT 105

RESULT 37

PCT-US94-07926-36
; Sequence 36, Application PC/TUS9407926
; GENERAL INFORMATION:
; APPLICANT: Tischfield, Jay A.
; APPLICANT: Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Mammalian Phospholipase A2 Nucleotide
; TITLE OF INVENTION: Sequences and Low Molecular Weight Amino Acid Sequences
; TITLE OF INVENTION: Encoded Thereby, Antisense Sequences and Nucleotide
; TITLE OF INVENTION: Sequences Having Internal Ribosome Binding Sites
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &
; ADDRESSEE: Russell PA
; STREET: 200 East Broward Boulevard
; CITY: Fort Lauderdale
; STATE: FL
; COUNTRY: USA
; ZIP: 33301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07926
; FILING DATE: 15-JUL-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/097,354
; FILING DATE: 26-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Manso, Peter J.
; REGISTRATION NUMBER: 32,264
; REFERENCE/DOCKET NUMBER: IN21044-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 305-527-2498
; TELEFAX: 305-764-4996
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 148 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-07926-36

Query Match 17.9%; Score 5; DB 5; Length 148;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GSAIT 6
|||||
Db 101 GSAIT 105

RESULT 38

US-08-679-493A-24
; Sequence 24, Application US/08679493A
; Patent No. 6303295
; GENERAL INFORMATION:
; APPLICANT: Taylor, Ethan W.
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS
; FILE REFERENCE: 55-95
; CURRENT APPLICATION NUMBER: US/08/679,493A
; CURRENT FILING DATE: 1996-07-12
; PRIOR APPLICATION NUMBER: 60/001203
; PRIOR FILING DATE: 1995-07-14
; PRIOR APPLICATION NUMBER: 60/003,112
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(158)
; OTHER INFORMATION: X is selenocysteine.
US-08-679-493A-24

Query Match 17.9%; Score 5; DB 4; Length 158;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LPSLC 17
|||||
Db 83 LPSLC 87

RESULT 39

US-08-360-606B-33
; Sequence 33, Application US/08360606B
; Patent No. 5919617
; GENERAL INFORMATION:
; APPLICANT: Jnanendra K. Bhattacharjee
; APPLICANT: Richard C. Garrad
; APPLICANT: Paul L. Skatrud
; APPLICANT: Robert P. Peery
; TITLE OF INVENTION: Methods and Reagents for
; TITLE OF INVENTION: Detecting Fungal Pathogens in a
; TITLE OF INVENTION: Biological Sample
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 S. Wacker Drive Suite 3200
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS Word 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,606B
; FILING DATE: December 21, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berghoff, Paul H.
; REGISTRATION NUMBER: 30,243
; REFERENCE/DOCKET NUMBER: 94,319
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)913-0001
; TELEFAX: (312)913-0002
; INFORMATION FOR SEQ ID NO: 33:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 161 amino acid residues
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: peptide
; HYPOTHETICAL: Yes
; FRAGMENT TYPE: Internal fragment
; ORIGINAL SOURCE:
; ORGANISM: Cryptococcus neoformans
US-08-360-606B-33

Query Match 17.9%; Score 5; DB 2; Length 161;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 AEENI 28
Db 131 AEENI 135

RESULT 40
US-08-804-180C-4
; Sequence 4, Application US/08804180C
; Patent No. 6107056
; GENERAL INFORMATION:
; APPLICANT: Martin K. Oaks
; TITLE OF INVENTION: sCTLA-4 and Its Soluble Products
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thomas M. Wozny
; STREET: 100 East Wisconsin Avenue
; CITY: Milwaukee
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.50 inch Disk
; COMPUTER: IBM
; OPERATING SYSTEM: DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804.180C
; FILING DATE: February 20, 1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Thomas M. Wozny
; REGISTRATION NUMBER: 28,922
; REFERENCE/DOCKET NUMBER: 3284-00003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 271-7590
; TELEFAX: (414) 271-5770
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: Polypeptide
; HYPOTHETICAL: no
; FRAGMENT TYPE: Internal fragment
; ORIGINAL SOURCE:
; ORGANISM: Rattus No. 6107056vegicus
; STRAIN: ACI
; DEVELOPMENTAL STAGE: Adult
; CELL TYPE: Splenocyte
; FEATURE:
; NAME/KEY: Rat sCTLA-4
; IDENTIFICATION METHOD: Found by experiment
; OTHER INFORMATION: B7 binding protein
US-08-804-180C-4

Query Match 17.9%; Score 5; DB 3; Length 174;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRA 24
Db 117 QGLRA 121

RESULT 41
US-08-067-684-14
; Sequence 14, Application US/08067684
; Patent No. 5434131
; GENERAL INFORMATION:
; APPLICANT: Linsley, Peter S.
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Damle, Nitin K.
; TITLE OF INVENTION: CTLA4 RECEPTOR AND METHODS FOR ITS USE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Suite 900
; CITY: Pasadena
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 91101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/067,684
; FILING DATE: 26-MAY-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Adriano, Sarah B.
; REGISTRATION NUMBER: 34,470
; REFERENCE/DOCKET NUMBER: 7848-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 310/312-9900
; TELEFAX: 310/479-8340
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-067-684-14

Query Match 17.9%; Score 5; DB 1; Length 187;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRA 24
Db 81 QGLRA 85

RESULT 42
US-08-008-898-14
; Sequence 14, Application US/08008898
; Patent No. 5770197
; GENERAL INFORMATION:
; APPLICANT: Linsley, Peter S.
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Damle, Nitin K.
; APPLICANT: Brady, William

; TITLE OF INVENTION: CTLA4 RECEPTOR AND METHODS FOR ITS USE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 201 South Lake Avenue, Suite 800
; CITY: Pasadena
; STATE: California
; COUNTRY: United States
; ZIP: 91101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/008,898
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/723,617
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandel, Saralynn
; REGISTRATION NUMBER: 31,853
; REFERENCE/DOCKET NUMBER: 7848
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-8321
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-008-898-14

Query Match 17.9%; Score 5; DB 1; Length 187;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0;

Qy 20 QGLRA 24
Db 81 QGLRA 85

RESULT 43
US-08-459-818-14
; Sequence 14, Application US/08459818
; Patent No. 5851795
; GENERAL INFORMATION:
; APPLICANT: Linsley, Peter S.
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Damle, Nitin K.
; APPLICANT: Brady, William
; TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 11150 Santa Monica Blvd., Suite 400
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FastSeq 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,818
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:
; NAME: Adriano, Sarah B.
; REGISTRATION NUMBER: 34,470
; REFERENCE/DOCKET NUMBER: 30436.35US02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 310-445-1140
; TELEFAX: 310-445-9031
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-459-818-14

Query Match 17.9%; Score 5; DB 2; Length 187;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0;

Qy 20 QGLRA 24
Db 81 QGLRA 85

RESULT 44
US-08-889-666-14
; Sequence 14, Application US/08889666
; Patent No. 5885579
; GENERAL INFORMATION:
; APPLICANT: Linsley, Peter S.
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Damle, Nitin K.
; APPLICANT: Brady, William
; APPLICANT: Kiener, Peter A.
; TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 11150 Santa Monica Blvd., Suite 400
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/889,666
; FILING DATE: 08-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/375390
; FILING DATE: 18-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Adriano, Sarah B.
; REGISTRATION NUMBER: 34,470
; REFERENCE/DOCKET NUMBER: 30436-35US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 310-445-1140
; TELEFAX: 310-445-9031
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-889-666-14

Query Match 17.9%; Score 5; DB 2; Length 187;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRA 24
|||||
Db 81 QGLRA 85

RESULT 45

US-08-465-078-14
; Sequence 14, Application US/08465078
; Patent No. 5885796
; GENERAL INFORMATION:
; APPLICANT: Linsley, Peter S.
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Damle, Nitin K.
; APPLICANT: Brady, William
; APPLICANT: Kiener, Peter A.
; TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 11150 Santa Monica Blvd., Suite 400
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,078
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/375390
; FILING DATE: 18-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Adriano, Sarah B.
; REGISTRATION NUMBER: 34,470
; REFERENCE/DOCKET NUMBER: 30436-35US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 310-445-1140
; TELEFAX: 310-445-9031
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-465-078-14

Query Match 17.9%; Score 5; DB 2; Length 187;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRA 24
|||||
Db 81 QGLRA 85

RESULT 46

US-08-725-776-14
; Sequence 14, Application US/08725776
; Patent No. 5968510
; GENERAL INFORMATION:
; APPLICANT: Linsley, Peter S.
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Damle, Nitin K.

; APPLICANT: Brady, William
; APPLICANT: Kiener, Peter A.
; TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 11150 Santa Monica Blvd., Suite 400
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/725,776
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/375390
; FILING DATE: 18-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Adriano, Sarah B.
; REGISTRATION NUMBER: 34,470
; REFERENCE/DOCKET NUMBER: 30436-35US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 310-445-1140
; TELEFAX: 310-445-9031
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-725-776-14

Query Match 17.9%; Score 5; DB 2; Length 187;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRA 24
|||||
Db 81 QGLRA 85

RESULT 47

US-08-488-062-14
; Sequence 14, Application US/08488062
; Patent No. 5977318
; GENERAL INFORMATION:
; APPLICANT: Linsley, Peter S.
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Damle, Nitin K.
; APPLICANT: Brady, William
; APPLICANT: Kiener, Peter A.
; TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 11150 Santa Monica Blvd., Suite 400
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/488,062
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/375390
; FILING DATE: 18-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Adriano, Sarah B.
; REGISTRATION NUMBER: 34,470
; REFERENCE/DOCKET NUMBER: 30436-35US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 310-445-1140
; TELEFAX: 310-445-9031
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-488-062-14

Query Match 17.9%; Score 5; DB 2; Length 187;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRA 24
Db 81 QGLRA 85

RESULT 48
US-08-228-208A-14
; Sequence 14, Application US/08228208A
; Patent No. 6090914
; GENERAL INFORMATION:
; APPLICANT: Linsley, Peter S.
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Damle, Nitin K.
; APPLICANT: Brady, William
; APPLICANT: Wallace, Philip M.
; TITLE OF INVENTION: CTLA4/CD28Ig HYBRID FUSION
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 11150 Santa Monica Boulevard, Suite 400
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/228,208A
; FILING DATE: 15-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/008,898
; FILING DATE: 22-JAN-1993
; APPLICATION NUMBER: 07/723,617
; FILING DATE: 27-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Adriano, Sarah B.
; REGISTRATION NUMBER: 34,470
; REFERENCE/DOCKET NUMBER: 30436-30US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 310 445-1140
; TELEFAX: 310 445-9031
; TELEX:

; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-228-208A-14

Query Match 17.9%; Score 5; DB 3; Length 187;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRA 24
Db 81 QGLRA 85

RESULT 49
PCT-US95-06726-36
; Sequence 36, Application PC/TUS9506726
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Ligands for Induction of Antigen Specific
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, suite 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/06726
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/253,783
; FILING DATE: 03 JUNE 1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Amy E. Mandragouras
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: RPI-016PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-06726-36

Query Match 17.9%; Score 5; DB 5; Length 187;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRA 24
Db 81 QGLRA 85

RESULT 50
US-08-882-704A-2
; Sequence 2, Application US/08882704A

```

; Patent No. 5879906
; GENERAL INFORMATION:
; APPLICANT: Jefferson, Richard A.
; APPLICANT: Wilson, Katherine J.
; APPLICANT: Leader, Michael
; TITLE OF INVENTION: GLUCURONIDE REPRESSORS AND USES THEREOF
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882.704A
; FILING DATE: 25-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5879906tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 190106.404
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 195 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-08-882-704A-2

```

```

Query Match      17.9%; Score 5; DB 2; Length 195;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 20 QGLRA 24
    |||||
Db 182 QGLRA 186

```

```

Search completed: August 22, 2002, 11:17:49
Job time: 56 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 11:18:19 ; Search time 27.33 Seconds
(without alignments)
98.445 Million cell updates/sec

Title: US-09-856-199-2

Perfect score: 28

Sequence: 1 TGSATYQCKENLPSLC SXOGLRAENI 28

Scoring table:

OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1200 summaries

Database :

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	11	39.3	380	2 T10442	selenoprotein P pr
2	11	39.3	381	1 A47327	selenoprotein P pr
3	8	28.6	385	1 OMRTSP	selenoprotein P pr
4	7	25.0	370	2 F84151	response regulator
5	6	21.4	71	2 AC0673	conserved hypotet
6	6	21.4	92	2 E90889	hypothetical prote
7	6	21.4	92	2 D85728	hypothetical prote
8	6	21.4	153	2 F95336	hypothetical prote
9	6	21.4	174	2 G83712	hypothetical prote
10	6	21.4	175	2 G69549	acetyl-CoA decarbo
11	6	21.4	186	2 S76419	probable ferredoxi
12	6	21.4	187	2 D83125	probable proteins
13	6	21.4	191	2 D90040	hypothetical prote
14	6	21.4	194	2 A86017	hypothetical prote
15	6	21.4	194	2 A91171	hypothetical prote
16	6	21.4	210	2 F81386	probable acidic pe
17	6	21.4	221	2 S39522	hypothetical prote
18	6	21.4	248	2 G86387	hypothetical prote
19	6	21.4	249	2 A97227	polyferredoxin [im
20	6	21.4	255	1 S39737	glucose 1-dehydrog
21	6	21.4	272	2 H95954	probable Sir2-like
22	6	21.4	276	2 AG2532	hypothetical prote
23	6	21.4	281	2 G24723	tryptophan synthas
24	6	21.4	297	2 S37636	hypothetical prote
25	6	21.4	301	2 J50681	hypothetical prote
26	6	21.4	304	2 S61033	hypothetical prote
27	6	21.4	324	1 B69553	methanol dehydroge
28	6	21.4	352	2 T20729	hypothetical prote
29	6	21.4	359	2 I50712	TOP AP - chicken

30	6	21.4	360	2 H71324	probable uracil ph
31	6	21.4	394	2 A45999	tyrosine--trna lig
32	6	21.4	394	2 S72814	hypothetical prote
33	6	21.4	405	2 A46699	translation initia
34	6	21.4	427	2 T23954	hypothetical prote
35	6	21.4	443	2 G83350	hypothetical prote
36	6	21.4	476	2 S36412	H+-transporting At
37	6	21.4	479	2 AE2117	hypothetical prote
38	6	21.4	480	2 T41971	alkaline exonuclea
39	6	21.4	480	2 S76598	hypothetical prote
40	6	21.4	481	2 S76820	hypothetical prote
41	6	21.4	485	2 G82676	ABC transporter me
42	6	21.4	486	2 T08907	probable ABC trans
43	6	21.4	487	1 S07062	glutamate receptor
44	6	21.4	487	2 S73161	hypothetical prote
45	6	21.4	489	2 C87480	conserved hypotet
46	6	21.4	489	2 AF2800	ABC transporter su
47	6	21.4	495	2 AB0703	conserved hypotet
48	6	21.4	500	2 B83364	pyoverdine biosynt
49	6	21.4	503	2 AD0293	conserved hypotet
50	6	21.4	507	2 AD3382	ABC transporter-as
51	6	21.4	508	2 C64926	ynhE protein - Esc
52	6	21.4	508	2 F90927	hypothetical prote
53	6	21.4	508	2 B85776	hypothetical prote
54	6	21.4	537	2 D96031	hypothetical prote
55	6	21.4	543	2 H97579	ABC transporter me
56	6	21.4	553	1 A42499	mullerian inhibiti
57	6	21.4	555	1 S20100	mullerian inhibiti
58	6	21.4	557	2 A85060	probable ABC trans
59	6	21.4	575	1 WFBOM	mullerian inhibiti
60	6	21.4	578	2 S25577	beta-hemolysin - A
61	6	21.4	663	2 D83545	probable helicase
62	6	21.4	684	2 S41788	endo-1,4-beta-xyla
63	6	21.4	702	2 S48753	major surface prot
64	6	21.4	718	1 ALBSMX	cyclomaltodextrin
65	6	21.4	819	2 T40527	hypothetical prote
66	6	21.4	896	2 T00389	hypothetical prote
67	6	21.4	1151	2 G96805	hypothetical prote
68	6	21.4	1214	2 G97419	streptococcal hema
69	6	21.4	1241	2 T18311	hypothetical prote
70	6	21.4	1248	2 A82637	conserved hypotet
71	6	21.4	1505	2 JC4851	hypoxia-inducible
72	6	21.4	1948	2 B69511	N conserved hypoth
73	5	17.9	15	2 S57201	basic proteinase I
74	5	17.9	35	2 JQ2193	hypothetical 3.9K
75	5	17.9	36	2 T14522	photosystem I prot
76	5	17.9	37	2 S31593	Ig heavy chain V r
77	5	17.9	40	2 S17921	photosystem I prot
78	5	17.9	64	2 S63681	signal transducer
79	5	17.9	73	2 D89782	hypothetical prote
80	5	17.9	82	2 G01978	calmodulin-depende
81	5	17.9	82	2 H90491	conserved hypotet
82	5	17.9	85	2 G85431	hypothetical prote
83	5	17.9	88	2 T03357	gene e10 protein -
84	5	17.9	89	2 S10002	C8-2 protein - hep
85	5	17.9	89	2 T14959	hypothetical prote
86	5	17.9	89	2 H82562	hypothetical prote
87	5	17.9	90	2 JCI163	homeotic protein 2
88	5	17.9	91	2 G01975	calmodulin-depende
89	5	17.9	92	2 D64901	hypothetical prote
90	5	17.9	93	2 A97726	hypothetical prote
91	5	17.9	95	2 JC7601	vesicular membrane
92	5	17.9	96	2 S22467	gene A9 protein pr
93	5	17.9	96	2 H97868	hypothetical prote
94	5	17.9	97	2 G70746	hypothetical prote
95	5	17.9	101	2 D97773	hypothetical prote
96	5	17.9	103	2 S20763	nonstructural prot
97	5	17.9	104	2 AD3608	hypothetical prote
98	5	17.9	105	2 G01976	calmodulin-depende
99	5	17.9	105	2 A42694	homeodomain protei
100	5	17.9	110	2 A82646	hypothetical prote
101	5	17.9	111	1 L2HUBO	Ig lambda chain V-
102	5	17.9	111	2 E75084	hypothetical prote

103	5	17.9	114	2	S50624	hypothetical prote	176	5	17.9	190	2	A89848	conserved hypothet
104	5	17.9	117	2	G71190	hypothetical prote	177	5	17.9	193	2	S25208	prsj protein - Esc
105	5	17.9	117	2	G90004	hypothetical prote	178	5	17.9	193	2	S16399	papj protein precu
106	5	17.9	118	2	D72609	hypothetical prote	179	5	17.9	193	2	T47408	lipid-transfer-lik
107	5	17.9	119	2	JN0736	hypothetical 13.7K	180	5	17.9	195	2	C97259	thiamine monophosp
108	5	17.9	120	2	P86824	hypothetical prote	181	5	17.9	195	2	T35879	hypothetical prote
109	5	17.9	121	2	G97521	YQZ2_BACSU protei	182	5	17.9	195	2	A97332	transcription regu
110	5	17.9	121	2	A12740	conserved hypothet	183	5	17.9	196	2	F90919	repressor for uid
111	5	17.9	122	2	AC1848	hypothetical prote	184	5	17.9	196	2	D64918	glucuronide repres
112	5	17.9	124	2	B83681	urease beta subuni	185	5	17.9	196	2	C85768	repressor for uid
113	5	17.9	126	2	T24650	hypothetical prote	186	5	17.9	199	2	A99218	15U ribosomal prot
114	5	17.9	130	2	T44467	transposase tnpB [187	5	17.9	199	2	E69826	probable 1'-acylgly
115	5	17.9	132	2	S31603	Ig heavy chain v r	188	5	17.9	200	2	PQ0250	polyprotein (clone
116	5	17.9	133	2	S14234	hypothetical prote	189	5	17.9	202	2	PQ0248	polyprotein (clone
117	5	17.9	135	2	A32317	sex-regulated prot	190	5	17.9	202	2	T23654	hypothetical prote
118	5	17.9	136	2	S76416	hypothetical prote	191	5	17.9	202	2	G97600	hypothetical prote
119	5	17.9	137	1	R5BY17	ribosomal protein	192	5	17.9	202	2	AH2822	conserved hypothet
120	5	17.9	138	2	A35676	Ig heavy chain pre	193	5	17.9	204	2	S75304	serine esterase -
121	5	17.9	139	2	T11639	ribosomal protein	194	5	17.9	206	2	H81932	probable dtmP kina
122	5	17.9	140	1	R5HU23	ribosomal protein	195	5	17.9	206	2	B81173	thymidylate kinase
123	5	17.9	140	1	R5RT23	ribosomal protein	196	5	17.9	206	2	E90090	60S ribosomal prot
124	5	17.9	140	2	JC1253	ribosomal protein	197	5	17.9	207	2	AC3555	hypothetical prote
125	5	17.9	140	2	P84744	60S ribosomal prot	198	5	17.9	208	2	A64089	hypothetical prote
126	5	17.9	140	2	T15337	hypothetical prote	199	5	17.9	208	2	G69901	acyl-carrier prote
127	5	17.9	140	2	T03693	ribosomal protein	200	5	17.9	209	2	T03532	precorrin isomeras
128	5	17.9	142	2	C64071	H+-transporting AT	201	5	17.9	209	2	H81382	hypothetical prote
129	5	17.9	144	2	T00215	type II secretion	202	5	17.9	209	2	AC3566	type I restriction
130	5	17.9	147	2	C95227	conserved hypothet	203	5	17.9	209	2	T22589	hypothetical prote
131	5	17.9	147	2	G98091	conserved hypothet	204	5	17.9	211	2	C70665	probable uref prot
132	5	17.9	147	2	H64380	hypothetical prote	205	5	17.9	212	1	R5HS1	ribosomal protein
133	5	17.9	148	1	PSHU	phospholipase A2 (206	5	17.9	212	2	G85357	hypothetical prote
134	5	17.9	149	2	B64438	hypothetical prote	207	5	17.9	212	2	E84741	probable synaptobr
135	5	17.9	150	2	R83018	hypothetical prote	208	5	17.9	213	2	B83179	hypothetical prote
136	5	17.9	151	2	S60870	phosphorylation-ac	209	5	17.9	214	2	C72372	pyrazinamidase/nic
137	5	17.9	152	2	AD0177	probable (3R)-hydr	210	5	17.9	217	2	T33353	hypothetical prote
138	5	17.9	153	2	S50431	hypothetical prote	211	5	17.9	218	2	A82343	hypothetical prote
139	5	17.9	154	2	P97805	hypothetical prote	212	5	17.9	218	2	D84478	hypothetical prote
140	5	17.9	157	2	B86177	hypothetical prote	213	5	17.9	218	2	D70012	conserved hypothet
141	5	17.9	157	2	T27453	hypothetical prote	214	5	17.9	219	1	A37243	hemopoietic cell s
142	5	17.9	158	2	B36152	ribosomal protein	215	5	17.9	219	2	AC2385	two-component resp
143	5	17.9	158	2	E86993	hypothetical prote	216	5	17.9	220	2	G69192	ABC transporter (g
144	5	17.9	159	2	B84669	hypothetical prote	217	5	17.9	220	2	F64203	glycerol uptake fa
145	5	17.9	160	2	PQ0254	polyprotein (clone	218	5	17.9	220	2	T46872	DNA repair protein
146	5	17.9	162	2	F64003	hypothetical prote	219	5	17.9	220	2	AH2299	hypothetical prote
147	5	17.9	162	2	A41503	AF/R1 pilus chain	220	5	17.9	221	2	F69526	conserved hypothet
148	5	17.9	162	2	A69543	hypothetical prote	221	5	17.9	221	2	A29063	cytotoxic T-lympho
149	5	17.9	164	2	T23249	hypothetical prote	222	5	17.9	223	2	T09536	cytotoxic T-lympho
150	5	17.9	166	2	JC1348	hypothetical prote	223	5	17.9	223	2	F64627	hypothetical prote
151	5	17.9	168	2	T45305	hypothetical 18K p	224	5	17.9	223	2	G71887	hypothetical prote
152	5	17.9	168	2	AE0192	hypothetical prote	225	5	17.9	224	2	B75460	hypothetical prote
153	5	17.9	173	2	PQ0251	probable C4-dicarb	226	5	17.9	226	2	AC0634	conserved hypothet
154	5	17.9	173	2	S24796	polypeptide (clone	227	5	17.9	227	2	D72508	hypothetical prote
155	5	17.9	175	2	PQ0253	coenzyme F420 hydr	228	5	17.9	228	2	C90557	n-acetylmanosamin
156	5	17.9	175	2	T06611	polypeptide (clone	229	5	17.9	229	2	A65114	hypothetical prote
157	5	17.9	178	1	IMEBAC	hypothetical prote	230	5	17.9	229	2	H91140	hypothetical prote
158	5	17.9	178	1	F90020	colicin A immunity	231	5	17.9	229	2	C85986	hypothetical prote
159	5	17.9	178	2	B82742	50S ribosomal prot	232	5	17.9	229	2	A95034	probable enzyme yh
160	5	17.9	182	2	PQ0252	hypothetical prote	233	5	17.9	230	2	AB0404	conserved hypothet
161	5	17.9	182	2	T44772	polypeptide (clone	234	5	17.9	231	2	S76365	L-ribulose-phospha
162	5	17.9	183	2	T37599	hypothetical prote	235	5	17.9	231	2	S55350	hypothetical prote
163	5	17.9	185	2	B97933	hypothetical prote	236	5	17.9	233	2	A69331	multicatalytic end
164	5	17.9	186	1	RPECTR	conserved hypothet	237	5	17.9	233	2	F80367	conserved hypothet
165	5	17.9	186	2	S08614	conjugal transfer	238	5	17.9	234	2	H98154	amino acid ABC tra
166	5	17.9	186	2	A95066	cytotoxic T-lympho	239	5	17.9	234	2	AD3133	conserved hypothet
167	5	17.9	187	1	A69973	acetyltransferase,	240	5	17.9	234	2	D84075	hypothetical prote
168	5	17.9	187	2	H84100	hypothetical prote	241	5	17.9	236	1	R3R22	ribosomal protein
169	5	17.9	187	2	AG1388	ribosomal protein	242	5	17.9	236	1	R3ZM2	ribosomal protein
170	5	17.9	187	2	A11763	conserved hypothet	243	5	17.9	236	1	R3WT2	ribosomal protein
171	5	17.9	188	2	R2351	conserved hypothet	244	5	17.9	236	2	S25746	Ig lambda chain -
172	5	17.9	189	1	R7SP12	hypothetical prote	245	5	17.9	236	2	S70156	hypothetical prote
173	5	17.9	189	2	T27837	ribosomal protein	246	5	17.9	236	2	T51661	myb-related transc
174	5	17.9	189	2	S68349	hypothetical prote	247	5	17.9	236	2	A97905	conserved hypothet
175	5	17.9	190	2	C71290	poli polypeptide -	248	5	17.9	237	2	T29759	hypothetical prote

249	5	17.9	238	2	I52638	neuronal growth-re	322	5	17.9	287	2	D70625	probable transpos
250	5	17.9	238	2	F84527	hypothetical prote	323	5	17.9	287	2	T01192	hypothetical prote
251	5	17.9	239	2	T39747	u2 snrp-specific	324	5	17.9	287	2	T48990	pirin-like protein
252	5	17.9	240	2	D75537	conserved hypot	325	5	17.9	288	1	TQEC13	Probable transpos
253	5	17.9	240	2	H90064	conserved hypot	326	5	17.9	289	2	AC1640	D-Amino Acid Amino
254	5	17.9	242	2	T50297	probable aminopept	327	5	17.9	289	2	E71939	hypothetical prote
255	5	17.9	244	2	T33929	dolichyl-phosphate	328	5	17.9	291	2	B64629	hypothetical prote
256	5	17.9	246	2	T17432	probable chaperone	329	5	17.9	292	2	E71885	hypothetical prote
257	5	17.9	246	2	AB0234	probable pilli asse	330	5	17.9	293	2	H90344	thiosulfate sulfur
258	5	17.9	246	2	T20341	hypothetical prote	331	5	17.9	294	2	F84922	hypothetical prote
259	5	17.9	248	2	AE1040	conserved hypot	332	5	17.9	295	2	D84902	probable WRKY-type
260	5	17.9	248	2	AE2908	conserved hypot	333	5	17.9	296	2	H83157	conserved hypot
261	5	17.9	250	2	S74495	hypothetical prote	334	5	17.9	296	2	E69025	conserved hypot
262	5	17.9	251	2	AG0132	probable hydroxyc	335	5	17.9	298	2	T40408	phosphoserine phos
263	5	17.9	253	2	AC8224	uridine phosphoryl	336	5	17.9	298	2	H83863	modulation of CheA
264	5	17.9	253	2	S05491	uridine phosphoryl	337	5	17.9	298	2	D86246	hypothetical prote
265	5	17.9	253	2	T46830	uridine phosphoryl	338	5	17.9	300	2	AE3255	3-oxoacyl-(acyl-ca
266	5	17.9	253	2	G86070	uridine phosphoryl	339	5	17.9	300	2	AE3255	ribosomal protein
267	5	17.9	253	2	A10460	uridine phosphoryl	340	5	17.9	300	2	AE2168	cation efflux syst
268	5	17.9	253	2	AF0916	uridine phosphoryl	341	5	17.9	301	2	S76240	hypothetical prote
269	5	17.9	253	2	T40687	hypothetical prote	342	5	17.9	303	2	S41303	yaid protein - Esc
270	5	17.9	254	2	AC2968	conserved hypot	343	5	17.9	303	2	D90684	recombination asso
271	5	17.9	255	2	D59448	conserved hypot	344	5	17.9	303	2	H85534	hypothetical prote
272	5	17.9	256	2	S34850	hypothetical prote	345	5	17.9	303	2	G96638	protein Tif9.22 [i
273	5	17.9	257	2	A96033	probable thiamine	346	5	17.9	305	2	I59376	translation initia
274	5	17.9	258	2	H98314	hypothetical prote	347	5	17.9	306	2	AB3049	arginase (imported
275	5	17.9	259	2	C64385	hypothetical prote	348	5	17.9	306	2	B84105	dipeptide ABC tran
276	5	17.9	260	2	AB2677	phosphoadenosine p	349	5	17.9	306	2	A48118	major epidermal ca
277	5	17.9	260	2	H97458	hypothetical prote	350	5	17.9	306	2	F69129	tungsten formylmet
278	5	17.9	260	2	T09497	myotubularin-relat	351	5	17.9	306	2	T11042	hypothetical prote
279	5	17.9	261	2	A95894	hypothetical arylm	352	5	17.9	307	2	D86838	hypothetical prote
280	5	17.9	263	2	AE1557	hypothetical prote	353	5	17.9	308	2	C71972	hypothetical prote
281	5	17.9	263	2	AG1199	hypothetical prote	354	5	17.9	310	2	S20889	sugar ABC transport
282	5	17.9	264	2	A64779	yhap protein - Esc	355	5	17.9	310	2	G96986	hypothetical prote
283	5	17.9	264	2	G90695	probable ligase [i	356	5	17.9	310	2	T37824	protein-tyrosine k
284	5	17.9	264	2	C85546	probable ligase yb	357	5	17.9	311	2	S13808	superantigen Mtv x
285	5	17.9	264	2	G97757	hypothetical prote	358	5	17.9	312	2	S34635	probable cobD - My
286	5	17.9	264	2	H95942	probable dimethyla	359	5	17.9	313	2	H70777	thymidylate synth
287	5	17.9	264	2	F97538	hypothetical prote	360	5	17.9	314	2	AB1309	thymidylate synth
288	5	17.9	264	2	A12757	conserved hypot	361	5	17.9	314	2	AB1681	probable ribosoma
289	5	17.9	268	2	B95150	Cof family protein	362	5	17.9	314	2	T43740	ribosomal protein
290	5	17.9	268	2	A98018	conserved hypot	363	5	17.9	314	2	AC1621	ribosomal protein
291	5	17.9	269	2	DB1344	hypothetical prote	364	5	17.9	314	2	AG1258	DNA-directed RNA p
292	5	17.9	271	2	S38046	hypothetical prote	365	5	17.9	315	2	T06944	stage V sporulatio
293	5	17.9	271	2	S72382	hypothetical prote	366	5	17.9	315	2	C83945	superantigen Mtv13
294	5	17.9	272	2	T15820	hypothetical prote	367	5	17.9	315	2	JH0554	superantigen Mtv13
295	5	17.9	272	2	C87479	acyltransferase fa	368	5	17.9	315	2	JH0551	superantigen Mtv1/
296	5	17.9	273	2	S48511	MP74 protein - yea	369	5	17.9	316	2	E75566	probable repressor
297	5	17.9	274	2	H72521	probable thiazole	370	5	17.9	317	2	A97511	hypothetical prote
298	5	17.9	274	2	T12791	hypothetical prote	371	5	17.9	317	2	AB2729	agmatinase [import
299	5	17.9	275	2	T41756	protein kinase PK1	372	5	17.9	317	2	B83390	probable transmem
300	5	17.9	275	2	H71690	hypothetical prote	373	5	17.9	317	2	G83287	hypothetical prote
301	5	17.9	276	2	E83832	ABC transporter (s	374	5	17.9	318	2	G97525	CG10208 protein (A
302	5	17.9	276	2	H96816	hypothetical prote	375	5	17.9	318	2	E97683	bili77_f3_136 prote
303	5	17.9	276	2	T47351	hypothetical prote	376	5	17.9	319	2	D70074	ABC transporter (b
304	5	17.9	277	2	T36371	hypothetical prote	377	5	17.9	320	2	E26795	superantigen Mtv(B
305	5	17.9	279	2	JC4327	uridine phosphoryl	378	5	17.9	320	2	S26174	superantigen Mtv17
306	5	17.9	279	2	E75444	branched-chain ami	379	5	17.9	320	2	S62193	methenyltetrahydro
307	5	17.9	279	2	AC2144	hypothetical prote	380	5	17.9	320	2	D69203	methenyltetrahydro
308	5	17.9	282	2	AG1534	hypothetical prote	381	5	17.9	321	2	F86864	homoserine O-succ
309	5	17.9	283	2	H64764	probable taurine d	382	5	17.9	321	2	AB5829	fucose synthetase
310	5	17.9	283	2	F90681	taurine dioxigenas	383	5	17.9	321	2	F90983	fucose synthetase
311	5	17.9	283	2	B85532	taurine dioxigenas	384	5	17.9	322	1	G70037	conserved hypot
312	5	17.9	283	2	AB2744	conserved hypot	385	5	17.9	322	2	G64694	arginase - Helicob
313	5	17.9	284	2	AB1354	hypothetical prote	386	5	17.9	322	2	F71808	arginase - Helicob
314	5	17.9	284	2	AC1724	hypothetical prote	387	5	17.9	322	2	S24574	superantigen Mtv(S
315	5	17.9	285	1	C70066	conserved hypot	388	5	17.9	322	2	A64023	hypothetical prote
316	5	17.9	285	2	B82842	spermidine synthas	389	5	17.9	324	1	QOMV6M	MTV superantigen -
317	5	17.9	285	2	E86521	polysaccharide hyd	390	5	17.9	324	2	S77342	probable pyruvate
318	5	17.9	285	2	A72101	polysaccharide hyd	391	5	17.9	324	2	D66834	transcription regu
319	5	17.9	286	2	G81264	hypothetical prote	392	5	17.9	325	2	S49019	superantigen - mou
320	5	17.9	286	2	S07533	puif 11/9A-2 prote	393	5	17.9	325	2	JC5276	HXC-26 protein - h
321	5	17.9	287	2	A27082	2,4-dichlorophenox	394	5	17.9	326	2	S23387	protein kinase (EC

395	5	17.9	327	2	D72503	probable modificat	468	5	17.9	370	2	AH2503	hypothetical prote
396	5	17.9	327	2	A98237	arginase (U39262)	469	5	17.9	370	2	AB2578	membrane-bound lyt
397	5	17.9	327	2	E64188	diptide transpor	470	5	17.9	371	2	A43830	alanine dehydrogen
398	5	17.9	327	2	F85708	unknown protein en	471	5	17.9	372	2	I52196	homeobox transcrip
399	5	17.9	329	2	D75190	histidinol-phospha	472	5	17.9	372	2	A46037	Hox-1.11 - mouse
400	5	17.9	329	2	G97944	hypothetical prote	473	5	17.9	373	1	H69281	conserved hypotet
401	5	17.9	330	2	T11180	probable N-acetyl-	474	5	17.9	373	1	T36889	probable iron-side
402	5	17.9	330	2	I40700	type II site-speci	475	5	17.9	375	2	T33439	hypothetical prote
403	5	17.9	330	2	T49644	hypothetical prote	476	5	17.9	375	2	I50589	transcription fact
404	5	17.9	332	2	JC5310	galactose represso	477	5	17.9	375	2	S35941	homeotic protein H
405	5	17.9	332	2	C95077	pneumococcal surfa	478	5	17.9	376	2	D96635	probable gibberell
406	5	17.9	332	2	T34928	hypothetical prote	479	5	17.9	376	2	C70399	S-adenosylmethioni
407	5	17.9	333	2	S35550	type II site-speci	480	5	17.9	376	2	E95879	probable sugar ABC
408	5	17.9	333	2	A31923	amalgam protein pr	481	5	17.9	377	1	MFNZ55	matrix protein - s
409	5	17.9	333	2	AD0266	oligopeptide trans	482	5	17.9	377	1	T16764	hypothetical prote
410	5	17.9	334	2	AC1544	serine phosphatase	483	5	17.9	377	2	A98315	hypothetical prote
411	5	17.9	334	2	AD1186	serine phosphatase	484	5	17.9	377	2	AB2968	conserved hypotet
412	5	17.9	335	2	D83232	hypothetical prote	485	5	17.9	379	2	B64473	hypothetical prote
413	5	17.9	336	2	S74462	hypothetical prote	486	5	17.9	379	2	D84742	hypothetical prote
414	5	17.9	336	2	T30504	hypothetical prote	487	5	17.9	381	2	E97406	mtLK protein (AF00
415	5	17.9	337	2	S74850	hypothetical prote	488	5	17.9	382	2	D97580	tyrosyl-tRNA synth
416	5	17.9	338	2	A64303	conserved hypotet	489	5	17.9	382	2	B61036	transforming growt
417	5	17.9	338	2	H72115	cationic amino aci	490	5	17.9	383	2	G71442	hypothetical prote
418	5	17.9	338	2	H86506	cationic amino aci	491	5	17.9	383	2	S76334	hypothetical prote
419	5	17.9	339	2	E64592	dihydroorotase - H	492	5	17.9	384	2	H90544	hypothetical prote
420	5	17.9	339	2	T36061	hypothetical prote	493	5	17.9	384	2	E82088	conserved hypotet
421	5	17.9	340	2	A36988	replication factor	494	5	17.9	385	2	H95413	probable integrase
422	5	17.9	340	2	S78536	tnpY protein - Clo	495	5	17.9	387	2	G64505	hypothetical prote
423	5	17.9	341	2	S66889	probable membrane	496	5	17.9	387	2	T27788	hypothetical prote
424	5	17.9	342	2	D95874	probable proline r	497	5	17.9	388	2	A29937	gastriclin (BC 3.4
425	5	17.9	343	2	AF2624	hypothetical prote	498	5	17.9	388	2	AF1383	internalin protein
426	5	17.9	344	2	T32869	hypothetical prote	499	5	17.9	389	2	T11611	probable X-Pro ami
427	5	17.9	344	2	JC5942	chemokine receptor	500	5	17.9	390	2	T38191	actin-like protein
428	5	17.9	344	2	C69537	hypothetical prote	501	5	17.9	390	2	S20392	phosphoprotein pho
429	5	17.9	345	2	AH0350	probable DNA-bind	502	5	17.9	390	2	B95954	probable membrane-
430	5	17.9	345	2	F95088	conserved hypotet	503	5	17.9	390	2	F81393	probable altronate
431	5	17.9	345	2	A97956	DNA polymerase III	504	5	17.9	391	2	AG3372	toluene tolerance
432	5	17.9	346	2	F75457	conserved hypotet	505	5	17.9	391	2	T35574	probable alanine r
433	5	17.9	347	2	D83335	transcription regu	506	5	17.9	392	2	S20225	actin-like protein
434	5	17.9	348	2	C96530	hypothetical prote	507	5	17.9	392	2	B75466	sensor histidine k
435	5	17.9	348	2	T24016	hypothetical prote	508	5	17.9	393	2	T50441	hypothetical prote
436	5	17.9	348	2	H90850	hypothetical prote	509	5	17.9	394	2	D65167	probable membrane
437	5	17.9	348	2	AI1889	hypothetical prote	510	5	17.9	394	2	D75256	queuine tRNA ribos
438	5	17.9	349	2	H85626	hypothetical prote	511	5	17.9	394	2	A75012	conserved hypotet
439	5	17.9	349	2	E90912	hypothetical prote	512	5	17.9	395	2	S20059	type II site-speci
440	5	17.9	349	2	B90819	hypothetical prote	513	5	17.9	395	2	AB2779	ABC transporter, m
441	5	17.9	350	2	T30379	hypothetical prote	514	5	17.9	396	1	OHCSAD	acyl-lacyl-carrier
442	5	17.9	351	2	AD1305	hypothetical prote	515	5	17.9	397	2	S52783	aspartic proteinas
443	5	17.9	352	2	G70620	probable argc prot	516	5	17.9	397	2	JC2524	phosphoprotein pho
444	5	17.9	352	2	T36332	hypothetical prote	517	5	17.9	397	2	T39507	kappa-carragenase
445	5	17.9	353	2	T09665	peroxidase (EC 1.1	518	5	17.9	397	2	AC1877	hypothetical prote
446	5	17.9	353	2	T30496	hypothetical prote	519	5	17.9	399	2	T19115	hypothetical prote
447	5	17.9	353	2	A43697	homeotic protein z	520	5	17.9	400	2	E82817	cytochrome p-450 h
448	5	17.9	355	2	H83193	conserved hypotet	521	5	17.9	400	2	A49147	bone morphogenetic
449	5	17.9	355	2	F69482	hypothetical prote	522	5	17.9	400	2	B64071	tyrosine-specific
450	5	17.9	355	2	T42541	hypothetical prote	523	5	17.9	401	2	D82429	acetyl-CoA acetyl
451	5	17.9	355	2	T15759	hypothetical prote	524	5	17.9	402	2	H82813	cytochrome p450-11
452	5	17.9	356	2	S71179	cinnamyl-alcohol d	525	5	17.9	402	2	T51791	ser/thr specific p
453	5	17.9	356	2	S50811	ser/thr specific p	526	5	17.9	402	2	T48935	acyl CoA reductase
454	5	17.9	356	2	T19015	hypothetical prote	527	5	17.9	402	2	H81296	probable integral
455	5	17.9	356	2	F97923	DNA-damage-inducib	528	5	17.9	403	2	G90955	tyrosine-specific
456	5	17.9	359	2	A96534	hypothetical prote	529	5	17.9	403	2	D85804	tyrosine-specific
457	5	17.9	360	2	T08581	cinnamyl-alcohol d	530	5	17.9	403	2	AD0748	tyrosine-specific
458	5	17.9	360	2	B71309	hypothetical prote	531	5	17.9	403	2	B69338	conserved hypotet
459	5	17.9	361	2	AD1398	hypothetical prote	532	5	17.9	405	2	T30324	type I site-specif
460	5	17.9	366	2	D83224	conserved hypotet	533	5	17.9	406	1	B71209	probable D-arabino
461	5	17.9	367	2	S72824	hypothetical prote	534	5	17.9	406	2	C75022	d-arabino 3-hexulo
462	5	17.9	368	2	C95019	hypothetical prote	535	5	17.9	407	2	AE3615	glycosyl transfera
463	5	17.9	368	2	T51200	hypothetical prote	536	5	17.9	408	2	B81417	hypothetical prote
464	5	17.9	369	1	A41147	protein-tyrosine-p	537	5	17.9	408	2	E86895	integral membrane
465	5	17.9	369	2	B75609	probable ABC trans	538	5	17.9	408	2	S58131	integral membrane
466	5	17.9	370	2	G87212	conserved hypotet	539	5	17.9	409	2	B85735	probable membrane
467	5	17.9	370	2	A97360	outer membrane lip	540	5	17.9	410	2	T18670	probable ubiquinol

541	5	17.9	411	2	T07806	acyl-[acyl-carrier	614	5	17.9	476	2	E69426	dihydropteroate sy
542	5	17.9	413	2	E82421	hypothetical prote	615	5	17.9	476	2	AG1051	probable transport
543	5	17.9	414	2	S59783	hypothetical prote	616	5	17.9	476	2	AG1031	hypothetical prote
544	5	17.9	414	2	A40350	transcription repr	617	5	17.9	476	2	T17330	hypothetical prote
545	5	17.9	415	2	A48273	delta/Y11/NF-E1/UC	618	5	17.9	478	2	D64895	probable membrane
546	5	17.9	415	2	AC3235	nitrotriacetate	619	5	17.9	478	2	G84956	protease do prec
547	5	17.9	415	2	F97558	hypothetical prote	620	5	17.9	479	2	D81423	probable iron-sulf
548	5	17.9	416	1	JC4952	transcription init	621	5	17.9	480	2	G64580	hypothetical prote
549	5	17.9	416	2	S50342	homeotic protein S	622	5	17.9	482	2	B98118	rhannulokinase (EC
550	5	17.9	416	2	AE2527	hypothetical prote	623	5	17.9	483	1	S77264	amidase slr0877 -
551	5	17.9	417	2	AB2801	tyrosyl-tRNA synth	624	5	17.9	484	2	T41527	probable trna nucl
552	5	17.9	418	2	S39833	hypothetical prote	625	5	17.9	484	2	D65230	hypothetical 52.9
553	5	17.9	419	2	D84372	pyruvate dehydroge	626	5	17.9	484	2	A86116	hypothetical prote
554	5	17.9	419	2	T23666	hypothetical prote	627	5	17.9	484	2	A98275	hypothetical prote
555	5	17.9	420	2	T42090	tryptophan synthas	628	5	17.9	484	2	C75609	amino acid ABC tra
556	5	17.9	420	2	T08691	hypothetical prote	629	5	17.9	486	2	D98284	argininosuccinate
557	5	17.9	421	2	T43534	transcription fact	630	5	17.9	486	2	E95165	conserved hypotet
558	5	17.9	421	2	E90883	hypothetical prote	631	5	17.9	486	2	D98031	conserved hypotet
559	5	17.9	425	2	S76760	transcription init	632	5	17.9	490	2	E70649	probable regulator
560	5	17.9	425	2	A84849	probable RING zinc	633	5	17.9	491	1	ACBOE	nicotinic acetylch
561	5	17.9	427	2	G98138	hypothetical prote	634	5	17.9	491	2	S14166	cyclin B2 - yeast
562	5	17.9	427	2	AD3149	glycosyltransferas	635	5	17.9	491	2	G64473	2-isopropylmalate
563	5	17.9	428	2	AD0675	membrane transport	636	5	17.9	493	1	ACRYE	nicotinic acetylch
564	5	17.9	429	2	S20050	transcription fact	637	5	17.9	493	1	ACMSE	nicotinic acetylch
565	5	17.9	429	2	T08562	hypothetical prote	638	5	17.9	493	2	S34775	nicotinic acetylch
566	5	17.9	430	2	H71476	probable rRNA meth	639	5	17.9	494	2	S15523	variant surface gl
567	5	17.9	431	2	D70433	5-enolpyruvylshik	640	5	17.9	496	2	F69196	2-isopropylmalate
568	5	17.9	431	2	C83858	5-enolpyruvylshik	641	5	17.9	496	2	G90433	ABC transporter, A
569	5	17.9	432	1	D64773	trigger factor [va	642	5	17.9	496	2	T41114	uroporphyrin methy
570	5	17.9	432	2	B90690	trigger factor [im	643	5	17.9	498	2	JC4727	mob protein A - Er
571	5	17.9	432	2	F85540	hypothetical prote	644	5	17.9	498	2	D84535	hypothetical prote
572	5	17.9	432	2	AB0558	trigger factor [im	645	5	17.9	499	2	A96514	hypothetical prote
573	5	17.9	433	2	S77340	hypothetical prote	646	5	17.9	501	2	E75004	thermostable carbo
574	5	17.9	434	2	AF0383	Trigger factor [im	647	5	17.9	501	2	S76563	hypothetical prote
575	5	17.9	435	2	S77156	processing protein	648	5	17.9	502	2	B46570	H+-transporting AT
576	5	17.9	437	2	JC5459	inulin fructotrans	649	5	17.9	506	2	T19787	hypothetical prote
577	5	17.9	437	2	S66069	conserved hypotet	650	5	17.9	506	2	G69546	hypothetical prote
578	5	17.9	441	2	A41591	endothelin recepto	651	5	17.9	513	1	S50216	translation initia
579	5	17.9	441	2	B82023	probable membrane-	652	5	17.9	513	2	T37612	cytosol aminopepti
580	5	17.9	441	2	C81244	membrane-bound lyt	653	5	17.9	515	2	C71158	probable thermosta
581	5	17.9	441	2	H82642	hypothetical prote	654	5	17.9	518	2	F75460	hypothetical prote
582	5	17.9	445	2	C83314	probable two-compo	655	5	17.9	518	2	A64130	hypothetical prote
583	5	17.9	445	2	T33617	hypothetical prote	656	5	17.9	518	2	AH3616	n-acetylmannosamin
584	5	17.9	446	2	A69750	erythromycin ester	657	5	17.9	519	2	D97487	BH2174 hypothetical
585	5	17.9	447	1	T43319	Delta5 fatty acid	658	5	17.9	519	2	AD2705	conserved hypotet
586	5	17.9	447	2	C82958	probable two-compo	659	5	17.9	520	2	T18124	probable capsid pr
587	5	17.9	448	1	D71960	DNA repair protein	660	5	17.9	523	1	A41648	protein-tyrosine-p
588	5	17.9	449	2	T48936	acyl CoA reductase	661	5	17.9	523	2	F90414	elongation factor
589	5	17.9	450	2	S38114	hypothetical prote	662	5	17.9	525	2	E86463	hypothetical prote
590	5	17.9	451	2	G82096	sodium-dependent t	663	5	17.9	526	2	F83268	probable FAD-depen
591	5	17.9	453	2	H86882	DNA repair protein	664	5	17.9	526	2	B70859	hypothetical prote
592	5	17.9	454	2	A82952	glucosamine-1-phos	665	5	17.9	529	2	S37545	H+-transporting AT
593	5	17.9	454	2	H88791	protein T13F2.1 [i	666	5	17.9	529	2	A95178	polysaccharide bio
594	5	17.9	456	1	G64547	DNA repair protein	667	5	17.9	529	2	S29976	hypothetical prote
595	5	17.9	457	2	T09932	probable phosphodi	668	5	17.9	529	2	JC7318	lectin MAL - Micro
596	5	17.9	458	1	E71033	tlld-related prote	669	5	17.9	530	2	S62439	hypothetical serin
597	5	17.9	458	2	C75180	tlld related prote	670	5	17.9	531	2	S09859	hypothetical prote
598	5	17.9	459	2	E86176	protein F19P19.10	671	5	17.9	533	2	A45392	RNA-directed RNA p
599	5	17.9	459	2	F83895	metal-tetracycline	672	5	17.9	534	2	T26238	hypothetical prote
600	5	17.9	464	2	S10171	3-isopropylmalate	673	5	17.9	536	2	S66716	glutamate--trNA li
601	5	17.9	465	2	G71931	probable outer mem	674	5	17.9	537	2	F70597	hypothetical prote
602	5	17.9	466	2	I39748	Na+-transporting A	675	5	17.9	540	2	F98044	polysaccharide tra
603	5	17.9	466	2	AE2999	argininosuccinate	676	5	17.9	541	2	S76017	hypothetical prote
604	5	17.9	466	2	AD0516	3-isopropylmalate	677	5	17.9	541	2	A84830	hypothetical prote
605	5	17.9	467	2	D95253	L-fucose kinase	678	5	17.9	543	2	S56830	probable purine nu
606	5	17.9	468	2	T21824	hypothetical prote	679	5	17.9	544	2	T36645	probable large int
607	5	17.9	469	2	B70428	general secretion	680	5	17.9	545	2	T02079	probable carbonate
608	5	17.9	470	2	T39921	hypothetical prote	681	5	17.9	545	2	H83342	periplasmic trehal
609	5	17.9	471	2	S14165	cyclin B1 - yeast	682	5	17.9	545	2	S46151	probable purine nu
610	5	17.9	471	2	AE3384	zinc metalloprotei	683	5	17.9	547	2	S15028	chromatin-binding
611	5	17.9	473	1	ERAD40	early E2A DNA-bind	684	5	17.9	547	2	B87384	sensory box protei
612	5	17.9	474	2	B83719	multidrug resistan	685	5	17.9	548	2	G70610	hypothetical prote
613	5	17.9	475	2	T39486	hypothetical prote	686	5	17.9	549	2	H70774	probable atpA prot

687	5	17.9	551	2	S50444	hypothetical prote	760	5	17.9	634	2	A97356	lon-like ATP-depen
688	5	17.9	553	2	T40438	hypothetical prote	761	5	17.9	635	2	T02536	CER1-like protein
689	5	17.9	554	2	A96018	probable regulator	762	5	17.9	635	2	C71021	hypothetical prote
690	5	17.9	556	1	S12602	60K cysteine-rich	763	5	17.9	640	2	G72339	threonine-tRNA li
691	5	17.9	556	2	JC5636	Ca2+/calmodulin-de	764	5	17.9	640	2	AG1269	threonyl-tRNA synt
692	5	17.9	556	2	A86560	60 kDa Cysteineri	765	5	17.9	640	2	AI1631	threonyl-tRNA synt
693	5	17.9	557	1	B39439	60K cysteine-rich	766	5	17.9	643	2	F06682	probable gamma-glu
694	5	17.9	557	2	T47506	hypothetical prote	767	5	17.9	645	2	F86627	LysR family transc
695	5	17.9	558	2	T09976	H+-transporting At	768	5	17.9	648	2	T20144	hypothetical prote
696	5	17.9	558	2	JC5204	60K cysteine-rich	769	5	17.9	649	2	B75317	threonyl-tRNA synt
697	5	17.9	560	1	WFHUM	mullerian inhibiti	770	5	17.9	650	2	T21498	hypothetical prote
698	5	17.9	561	2	T11378	NADH dehydrogenase	771	5	17.9	651	2	PC1123	hypothetical prote
699	5	17.9	562	2	C81233	long-chain-fatty-a	772	5	17.9	652	2	T03504	probable DNA topoi
700	5	17.9	564	2	C42523	A55K protein - vac	773	5	17.9	652	2	D85044	hypothetical prote
701	5	17.9	564	2	JQ1792	SalF17R protein -	774	5	17.9	653	2	T02080	probable carbonate
702	5	17.9	564	2	H83397	probable two-compo	775	5	17.9	653	2	B90908	probable host spec
703	5	17.9	564	2	T49322	related to RNA-bin	776	5	17.9	656	2	A84018	hypothetical prote
704	5	17.9	564	2	AH2321	hypothetical prote	777	5	17.9	659	2	S38108	hypothetical prote
705	5	17.9	565	2	A10479	probable membrane	778	5	17.9	664	2	T20420	hypothetical prote
706	5	17.9	565	2	S75255	tyrs protein sir10	779	5	17.9	664	2	T29011	hypothetical prote
707	5	17.9	565	2	T47423	hypothetical prote	780	5	17.9	666	1	D69103	DNA helicase (EC 3
708	5	17.9	566	2	AF2116	hypothetical prote	781	5	17.9	666	1	B70803	hypothetical prote
709	5	17.9	567	2	A87394	acyl-CoA synthetas	782	5	17.9	670	2	T05495	hypothetical prote
710	5	17.9	568	2	T49962	hypothetical prote	783	5	17.9	672	2	A2076	hypothetical prote
711	5	17.9	569	2	S64957	asparillopepsin I	784	5	17.9	672	2	B86548	hypothetical prote
712	5	17.9	573	2	S29623	hydroxymethylgluta	785	5	17.9	676	1	WMBEX6	UL6 protein - huma
713	5	17.9	578	2	B37852	phosphotransferase	786	5	17.9	678	2	H88187	protein C18H9.8 [i
714	5	17.9	578	2	A50428	probable exported	787	5	17.9	681	2	F83375	hypothetical prote
715	5	17.9	582	2	H70426	neutral proteinase	788	5	17.9	685	1	A48289	neurotrophic recep
716	5	17.9	586	2	D69250	RNase L inhibitor	789	5	17.9	685	2	E82297	c-di-GMP phosphodi
717	5	17.9	586	2	D82484	Sgat protein VCA02	790	5	17.9	687	2	S43250	thimet oligopeptid
718	5	17.9	587	2	T41173	phosphomannomutase	791	5	17.9	687	2	T33077	hypothetical prote
719	5	17.9	587	2	AC1510	internalin protein	792	5	17.9	694	2	D81280	hypothetical prote
720	5	17.9	589	2	AB1151	internalin protein	793	5	17.9	694	2	TFRBP	folliotropin recepr
721	5	17.9	590	2	S66956	hypothetical prote	794	5	17.9	696	2	JC7361	transferrin precur
722	5	17.9	591	2	T14364	probable transcrip	795	5	17.9	696	2	T50457	polysphosphate kina
723	5	17.9	593	2	A45135	profilaggrin - hum	796	5	17.9	699	2	T50457	folliotropin recepr
724	5	17.9	593	2	A96783	unknown protein F2	797	5	17.9	706	1	A45416	protein kinase [im
725	5	17.9	593	2	JC4884	organic cation tra	798	5	17.9	706	1	A45416	protein kinase C (
726	5	17.9	593	2	AI1235	internalin protein	799	5	17.9	707	1	A44500	protein kinase C (
727	5	17.9	594	2	T03922	hypothetical prote	800	5	17.9	715	2	B71418	hypothetical prote
728	5	17.9	597	2	JQ0107	hypothetical 66K p	801	5	17.9	722	2	C88397	hypothetical prote
729	5	17.9	597	2	B82140	C4-dicarboxylate t	802	5	17.9	722	2	H96986	protein H04J21.3 [
730	5	17.9	598	2	C82194	ATP-dependent prot	803	5	17.9	724	2	T00495	endo-1,4-beta gluc
731	5	17.9	603	2	S76615	hypothetical prote	804	5	17.9	724	2	F80187	L5 protein - white
732	5	17.9	603	2	D87438	metalloproteinase M	805	5	17.9	728	2	F50719	proteinase II [imp
733	5	17.9	603	2	T46236	hypothetical prote	806	5	17.9	728	2	T87187	C-Delta-1 - chicke
734	5	17.9	603	2	T37518	probable transcrip	807	5	17.9	729	2	T45780	sugar transporter-
735	5	17.9	605	2	A70030	hypothetical prote	808	5	17.9	732	2	T08420	1-phosphatidylinos
736	5	17.9	606	2	A33496	sensory kinase (EC	809	5	17.9	732	2	F91168	zinc-transporting
737	5	17.9	606	2	G86449	F5b14.23 protein -	810	5	17.9	732	2	F86014	zinc-transporting
738	5	17.9	608	2	T14731	UDPglucose--starch	811	5	17.9	732	2	S47688	lead, cadmium, zin
739	5	17.9	609	2	T04946	hypothetical prote	812	5	17.9	733	2	AB0991	heavy metal-transp
740	5	17.9	610	2	T41399	probable cyclophil	813	5	17.9	733	2	T24977	GGDEF family prote
741	5	17.9	611	1	W1LEP	El protein - Europ	814	5	17.9	735	2	F87355	transcription fact
742	5	17.9	611	2	AI0864	exonuclease V alph	815	5	17.9	736	2	T00023	TonB-dependent rec
743	5	17.9	614	1	B70772	probable 3'-phosph	816	5	17.9	737	2	D87350	hypothetical prote
744	5	17.9	617	1	MBBYO	mitochondrial oute	817	5	17.9	739	2	T15215	hypothetical prote
745	5	17.9	618	2	T09153	glucose-6-phosphat	818	5	17.9	747	2	S37694	gene PC326 protein
746	5	17.9	620	2	JC4925	Xmni methyltransfe	819	5	17.9	749	2	T50397	probable serine/th
747	5	17.9	621	2	JC1346	dopamine beta-mono	820	5	17.9	751	2	AD0345	probable phosphate
748	5	17.9	622	2	S75452	hypothetical prote	821	5	17.9	751	2	D72338	(P)pppp synthetas
749	5	17.9	623	2	G90690	probable proteinase	822	5	17.9	751	2	AG1329	penicillin-binding
750	5	17.9	623	2	C85541	probable proteinase	823	5	17.9	752	2	AG1700	hypothetical prote
751	5	17.9	626	2	S59435	hypothetical prote	824	5	17.9	755	2	T46411	hypothetical prote
752	5	17.9	627	2	T02197	hypothetical prote	825	5	17.9	758	2	T15577	hypothetical prote
753	5	17.9	629	2	G84481	probable receptor-	826	5	17.9	760	2	S64023	Alk1 protein - yea
754	5	17.9	629	2	G83251	probable acetyltra	827	5	17.9	760	2	D83743	xanthine dehydroge
755	5	17.9	632	2	T22888	hypothetical prote	828	5	17.9	764	2	D81883	probable outab-mem
756	5	17.9	632	2	AG1366	phosphotransferase	829	5	17.9	764	2	T39878	hypothetical prote
757	5	17.9	632	2	AH1735	phosphotransferase	830	5	17.9	766	2	AF2279	hypothetical prote
758	5	17.9	633	2	T41332	casp homolog - fis	831	5	17.9	767	2	E96813	probable Oxidodqua
759	5	17.9	633	2	D75112	glu-tRNA amidotran	832	5	17.9	770	2	T09026	hypothetical prote
												G90996	probable host spec

833	5	17.9	776	2	T33543	hypothetical prote	906	5	17.9	932	2	T49235	kinesin-like prote
834	5	17.9	781	1	Q0BEN7	helicase (EC 3.6.1	907	5	17.9	935	2	S57080	hypothetical prote
835	5	17.9	781	2	T42959	helicase (EC 3.6.1	908	5	17.9	936	2	F75622	hypothetical prote
836	5	17.9	781	2	H84501	En/Spm-like transp	909	5	17.9	938	2	T25215	hypothetical prote
837	5	17.9	781	2	A37956	sulfate permease I	910	5	17.9	954	2	T02657	hypothetical prote
838	5	17.9	782	2	S55094	probable membrane	911	5	17.9	956	2	JH0826	glutamate ionotrop
839	5	17.9	782	2	A85693	hypothetical prote	912	5	17.9	956	2	J50685	glutamate receptor
840	5	17.9	783	2	T00354	hypothetical prote	913	5	17.9	965	2	S58393	PET309 protein - y
841	5	17.9	785	2	C88131	protein F1067.1 [i	914	5	17.9	968	2	I60107	alanine--trNA liga
842	5	17.9	785	2	S54016	SOK2 protein - yea	915	5	17.9	970	2	A13605	potassium efflux s
843	5	17.9	787	2	A75347	GTP pyrophosphokin	916	5	17.9	973	2	A75135	probable DEAH ATP
844	5	17.9	788	1	Q0BEE3	HHLFI protein - hu	917	5	17.9	976	2	T29583	hypothetical prote
845	5	17.9	789	2	B87461	polyposphatase kina	918	5	17.9	977	2	I52657	seizure-related pr
846	5	17.9	790	2	S61587	transcription acti	919	5	17.9	980	2	T05414	protein kinase hom
847	5	17.9	794	2	E95942	probable aldehyde	920	5	17.9	980	2	T14756	hypothetical prote
848	5	17.9	797	2	T22294	hypothetical prote	921	5	17.9	991	2	A96604	hypothetical prote
849	5	17.9	801	1	D70309	ribonucleoside-dip	922	5	17.9	991	2	H86168	hypothetical prote
850	5	17.9	809	2	T20430	hypothetical prote	923	5	17.9	1002	1	GNWXS	genome polyprotein
851	5	17.9	810	2	H88565	protein ZK632.1 [i	924	5	17.9	1004	2	T38074	hypothetical prote
852	5	17.9	811	2	E96791	hypothetical prote	925	5	17.9	1007	2	S48535	rho-type GTPase-ac
853	5	17.9	812	2	T34180	hypothetical prote	926	5	17.9	1010	2	T34314	hypothetical prote
854	5	17.9	813	2	T31214	hypothetical prote	927	5	17.9	1010	2	C81524	glycyl--trNA synthe
855	5	17.9	813	2	T04313	protein kinase Xa2	928	5	17.9	1014	2	AC1971	glycyl trNA synthe
856	5	17.9	814	2	T47641	hypothetical prote	929	5	17.9	1014	2	H86608	hypothetical prote
857	5	17.9	817	2	S53919	hypothetical prote	930	5	17.9	1014	2	T24412	hypothetical prote
858	5	17.9	819	2	B87580	conserved hypotet	931	5	17.9	1014	2	T31433	Na+/Ca2+ K+-exchan
859	5	17.9	822	2	C71633	ATP-dependent nucl	932	5	17.9	1014	2	A72016	glycine--trNA liga
860	5	17.9	825	2	JC4103	DNA-binding protei	933	5	17.9	1018	2	E83099	probable RND efflu
861	5	17.9	827	2	S25949	S-layer protein -	934	5	17.9	1023	2	JC4013	major acidic nucle
862	5	17.9	827	2	S25949	gene coxII intron	935	5	17.9	1024	2	T30868	RhoA-binding prote
863	5	17.9	830	1	A44047	glycoprotein B pre	936	5	17.9	1026	2	S73460	hypothetical prote
864	5	17.9	830	1	B44047	glycoprotein B pre	937	5	17.9	1030	2	G87346	probable helicase
865	5	17.9	830	2	T43959	glycoprotein B [im	938	5	17.9	1032	1	G78PT4	baseplate protein
866	5	17.9	830	2	T44186	probable glycoprot	939	5	17.9	1035	2	T42093	phospholipase D (E
867	5	17.9	831	1	VGBES5	glycoprotein B - h	940	5	17.9	1036	2	T13732	phospholipase D (E
868	5	17.9	832	2	G86147	protein T1N6.6 [im	941	5	17.9	1037	2	T13943	phospholipase D (E
869	5	17.9	834	2	T00512	serine/threonine-s	942	5	17.9	1037	2	E84723	hypothetical prote
870	5	17.9	838	2	JC7363	95K retinoblastoma	943	5	17.9	1042	2	E95103	DNA polymerase III
871	5	17.9	840	2	T36829	probable phenylala	944	5	17.9	1042	2	C97971	DNA-directed DNA p
872	5	17.9	840	2	T33217	hypothetical prote	945	5	17.9	1044	2	S16516	integrin alpha-8 c
873	5	17.9	842	2	S49124	carbon-monoxide de	946	5	17.9	1051	2	S55259	TIF1 protein - mou
874	5	17.9	846	1	Q0BEC3	HQRFL protein - hu	947	5	17.9	1052	2	AF2959	conserved hypotet
875	5	17.9	847	2	F75270	cation-transportin	948	5	17.9	1056	2	A53767	mucin MUC5B, trach
876	5	17.9	851	2	D90216	hypothetical prote	949	5	17.9	1058	2	T30178	mitotic checkpoint
877	5	17.9	863	2	T47038	hypothetical prote	950	5	17.9	1060	2	S63252	hypothetical prote
878	5	17.9	863	2	AD0234	probable fibrillar	951	5	17.9	1067	2	T28663	probable phage hos
879	5	17.9	865	2	A53186	flug protein - Eme	952	5	17.9	1067	2	AE0260	hypothetical prote
880	5	17.9	878	1	A40091	interleukin-3 rece	953	5	17.9	1074	2	T17203	phospholipase (EC
881	5	17.9	880	2	T38083	potassium transpor	954	5	17.9	1074	2	T13725	phospholipase D (E
882	5	17.9	880	2	F75103	conserved hypotet	955	5	17.9	1075	2	T46635	phospholipase D (E
883	5	17.9	880	2	F83386	hypothetical prote	956	5	17.9	1080	2	T03964	probable ubiquitin
884	5	17.9	880	2	D69427	conserved hypotet	957	5	17.9	1083	2	T05689	hypothetical prote
885	5	17.9	884	2	C70729	hypothetical prote	958	5	17.9	1085	2	T18369	K-Cl cotransport p
886	5	17.9	885	2	S59660	anaphase spindle e	959	5	17.9	1089	2	T36663	protein kinase, tr
887	5	17.9	889	2	T30715	probable major cor	960	5	17.9	1094	2	A53435	vesicular transpor
888	5	17.9	892	2	AE0650	alcohol dehydrogen	961	5	17.9	1095	2	T01916	hypothetical prote
889	5	17.9	896	1	A35782	cytokine receptor	962	5	17.9	1095	2	PC1114	SKCDC25 protein -
890	5	17.9	896	2	F81994	probable [glutamat	963	5	17.9	1102	2	T31004	probable mitotic c
891	5	17.9	896	2	B81223	glutamate-ammonia-	964	5	17.9	1122	2	T18346	MG1 protein precu
892	5	17.9	900	2	B87252	sensory box histid	965	5	17.9	1125	2	T14892	transcription fact
893	5	17.9	900	2	A95340	cation transport p	966	5	17.9	1127	2	D70671	pyruvate carboxyla
894	5	17.9	909	1	A54809	disease resistance	967	5	17.9	1132	2	C75259	probable iron-sulf
895	5	17.9	910	2	E96770	hypothetical prote	968	5	17.9	1147	2	D87295	smc protein [impor
896	5	17.9	912	2	H90567	hypothetical prote	969	5	17.9	1154	2	T18525	diacylglycerol kin
897	5	17.9	919	2	T05746	hypothetical prote	970	5	17.9	1154	2	A13431	chromosome segrega
898	5	17.9	921	2	S40495	collagen alpha 1(I	971	5	17.9	1155	2	AC2675	chromosome segrega
899	5	17.9	921	2	S42617	collagen alpha 1(I	972	5	17.9	1156	2	T34852	probable secreted
900	5	17.9	921	2	A48184	transcription init	973	5	17.9	1159	2	E70741	probable regulator
901	5	17.9	921	2	A45183	TAp-associated fac	974	5	17.9	1165	2	A97457	structural mainten
902	5	17.9	925	2	JC2033	G protein-coupled	975	5	17.9	1166	1	S06142	protein-tyrosine k
903	5	17.9	926	2	AH1034	hypothetical prote	976	5	17.9	1174	2	T08196	hypothetical prote
904	5	17.9	929	2	T28927	hypothetical prote	977	5	17.9	1175	2	T20346	pyruvate carboxyla
905	5	17.9	930	2	T42241	myotubularin prote	978	5	17.9	1198	2	B88279	protein C08H9.2 [i

979	5	17.9	1212	2	A96971	cobalamine-depende	1052	5	17.9	2101	2	A42184	nuclear mitotic ap
980	5	17.9	1220	2	T19117	hypothetical prote	1053	5	17.9	2116	2	T49818	glutamate synthase
981	5	17.9	1221	2	E83327	conserved hypotet	1054	5	17.9	2140	2	T18543	probable cell-adhe
982	5	17.9	1231	2	T35227	probable nitrate r	1055	5	17.9	2142	2	B35098	MHC class III hist
983	5	17.9	1258	2	T75453	5-methyltetrahydro	1056	5	17.9	2142	2	D86303	Fl7F16.1 protein -
984	5	17.9	1260	2	T14276	myosin-like protei	1057	5	17.9	2163	2	S50675	pre-mRNA splicing
985	5	17.9	1270	2	T01334	hypothetical prote	1058	5	17.9	2383	2	D64962	probable membrane
986	5	17.9	1285	2	AD0332	hypothetical prote	1059	5	17.9	2403	2	T30875	PRP8 protein homol
987	5	17.9	1276	2	T27859	hypothetical prote	1060	5	17.9	2475	2	S35307	polyprotein pp220
988	5	17.9	1296	1	B7C1AB	bontoxilysin (EC 3	1061	5	17.9	2475	2	B53435	vesicular transport
989	5	17.9	1296	2	I40645	bontulinum neurotox	1062	5	17.9	2550	2	T18552	saframycin Mxl syn
990	5	17.9	1298	2	T47523	DNA-binding protei	1063	5	17.9	2607	2	T31678	bacitracin synthet
991	5	17.9	1306	2	T13592	hypothetical prote	1064	5	17.9	2660	2	E85822	probable invasiv
992	5	17.9	1313	1	VGVUPT	M polyprotein - Pu	1065	5	17.9	2677	2	A38194	desmoplakin I - hu
993	5	17.9	1315	2	T50262	probable nucleopor	1066	5	17.9	2700	2	D88450	protein F21H11.2 l
994	5	17.9	1324	2	T14070	peptide synthetase	1067	5	17.9	2723	2	T03221	probable polyketid
995	5	17.9	1324	2	T13123	DNA replication pr	1068	5	17.9	2748	2	S57976	nuclear migration
996	5	17.9	1333	2	AF2020	two-component hybr	1069	5	17.9	3010	1	GNWVTC	genome polyprotein
997	5	17.9	1337	2	A53824	nuclear pore membr	1070	5	17.9	3010	1	A45573	genome polyprotein
998	5	17.9	1339	2	T40245	probable transcrip	1071	5	17.9	3010	1	S18030	genome polyprotein
999	5	17.9	1341	2	H98323	hypothetical prote	1072	5	17.9	3010	1	GNWVTC	genome polyprotein
1000	5	17.9	1345	2	T41960	major capsid prote	1073	5	17.9	3034	2	T14119	seven-pass transme
1001	5	17.9	1345	2	H90975	hypothetical prote	1074	5	17.9	3079	1	RGBY12	probable Gnpase-ac
1002	5	17.9	1350	2	T30341	zinc finger protei	1075	5	17.9	3172	2	S22012	erythronolide synt
1003	5	17.9	1355	2	T32092	hypothetical prote	1076	5	17.9	3178	2	S13595	6-deoxyerythronol
1004	5	17.9	1366	2	T42654	hypothetical prote	1077	5	17.9	3413	2	T17467	rifamycin polyketi
1005	5	17.9	1369	2	S58160	hypothetical prote	1078	5	17.9	3507	2	T34513	hypothetical prote
1006	5	17.9	1371	2	T29019	hypothetical prote	1079	5	17.9	3705	2	AD0123	probable autotrans
1007	5	17.9	1407	1	S28589	trichohyalin - rab	1080	5	17.9	3738	2	T05501	hypothetical prote
1008	5	17.9	1415	2	C83070	conserved hypotet	1081	5	17.9	3788	2	T13960	beige protein homo
1009	5	17.9	1420	2	B57062	SKB9 protein - yea	1082	5	17.9	4725	1	A44357	dynamin heavy chain
1010	5	17.9	1426	2	T03037	hypothetical prote	1083	5	17.9	4957	2	T03455	ALR protein - huma
1011	5	17.9	1463	2	A36861	orf 1b protein - L	1084	5	17.9	5147	1	IJFFMT	cadherin-related t
1012	5	17.9	1465	2	D86478	protein F1504.11 l	1085	5	17.9	5262	2	T03454	ALR protein - huma
1013	5	17.9	1488	2	C70984	probable ppsE prot	1086	5	17.9	7576	2	T17428	PK506 polyketide s
1014	5	17.9	1494	2	T14355	protein-cytosine-p	1087	5	17.9	13055	2	T16580	hypothetical prote
1015	5	17.9	1510	2	T13634	probable minor tal	1088	4	14.3	8	2	I48934	apolipoprotein A-I
1016	5	17.9	1525	2	T14961	hypothetical prote	1089	4	14.3	17	2	A39111	1g light chain - p
1017	5	17.9	1526	2	T19473	hypothetical prote	1090	4	14.3	19	2	I53673	amylid protein -
1018	5	17.9	1528	2	T14279	myosin-like protei	1091	4	14.3	20	2	A60525	lysosome (EC 3.2.1
1019	5	17.9	1539	2	G70630	probable cupH prot	1092	4	14.3	20	2	A57106	hull allergen Gly
1020	5	17.9	1546	2	G90503	lipoprotein [impor	1093	4	14.3	26	2	S68644	nicotinic acetylch
1021	5	17.9	1548	2	S54723	UDP-glucose-glyco	1094	4	14.3	27	2	B41575	bombinin-like pept
1022	5	17.9	1593	2	T22028	hypothetical prote	1095	4	14.3	28	2	T14210	NADH dehydrogenase
1023	5	17.9	1600	2	AB3281	NAD-specific gluta	1096	4	14.3	28	2	S38524	rRNA N-glycosidase
1024	5	17.9	1606	2	T34073	paranemin - chicke	1097	4	14.3	28	2	T14905	hypothetical prote
1025	5	17.9	1607	2	T43212	insulin-like growt	1098	4	14.3	28	2	S64701	hypothetical prote
1026	5	17.9	1620	2	S21045	complement protein	1099	4	14.3	30	2	S08565	ribulose-bisphosph
1027	5	17.9	1640	2	T09522	clathrin heavy cha	1100	4	14.3	31	2	A44926	flagellin 35K isot
1028	5	17.9	1663	1	C3HU	complement C3 prec	1101	4	14.3	33	2	I52083	major acute phase
1029	5	17.9	1664	2	T18262	S-layer protein -	1102	4	14.3	33	2	PQ0305	glycoprotein B - h
1030	5	17.9	1687	2	T30176	EGF repeat transme	1103	4	14.3	33	2	H82820	hypothetical prote
1031	5	17.9	1711	2	AB1283	peptidoglycan link	1104	4	14.3	34	2	T06870	photosystem I chai
1032	5	17.9	1732	2	G84664	hypothetical prote	1105	4	14.3	35	2	F81179	photosystem I prot
1033	5	17.9	1749	2	S75138	hypothetical prote	1106	4	14.3	35	2	T11812	photosystem I prot
1034	5	17.9	1753	2	T00350	hypothetical prote	1107	4	14.3	35	2	E81582	photosystem I 4K p
1035	5	17.9	1760	2	E86201	protein F12K11.4 l	1108	4	14.3	36	1	ALR21	photosystem I 4K p
1036	5	17.9	1765	2	T42388	sodium channel alp	1109	4	14.3	36	2	A34302	photosystem I prot
1037	5	17.9	1783	2	AC2091	serine/threonine k	1110	4	14.3	36	2	S17320	photosystem I prot
1038	5	17.9	1786	1	H71527	probable exincule	1111	4	14.3	36	2	S17324	photosystem I prot
1039	5	17.9	1787	2	F84528	probable retroelem	1112	4	14.3	36	2	S21985	photosystem I prot
1040	5	17.9	1799	2	AD1895	serine/threonine k	1113	4	14.3	36	2	S58561	photosystem I prot
1041	5	17.9	1823	2	S28974	vitellogenin precu	1114	4	14.3	36	2	T14832	photosystem I prot
1042	5	17.9	1840	2	T29091	transitin - chicke	1115	4	14.3	37	2	A85744	unknown protein en
1043	5	17.9	1841	2	T38091	cell division cont	1116	4	14.3	40	2	D82440	hypothetical prote
1044	5	17.9	1870	2	S37671	MHC class III hist	1117	4	14.3	41	2	C87355	hypothetical prote
1045	5	17.9	1872	2	S36152	MHC class III hist	1118	4	14.3	43	2	T07124	RNA-directed DNA p
1046	5	17.9	1898	1	A45973	trichohyalin - hum	1119	4	14.3	44	2	T44575	hypothetical prote
1047	5	17.9	1974	2	T16703	hypothetical prote	1120	4	14.3	44	2	S56313	GURF-2a protein -
1048	5	17.9	2052	2	T18519	myosin X - bovine	1121	4	14.3	45	2	E95058	hypothetical prote
1049	5	17.9	2054	2	T32413	probable acetyl-Co	1122	4	14.3	45	2	B81209	hypothetical prote
1050	5	17.9	2076	2	T28915	hypothetical prote	1123	4	14.3	45	2	F69384	hypothetical prote
1051	5	17.9	2089	2	C85426	AMF-like protein l	1124	4	14.3	45	2	D55679	cannabinoid recept

1125 4 14.3 46 1 VCBPPF coat protein - pha
1126 4 14.3 46 2 C81372 very hypothetical
1127 4 14.3 46 2 E99830 hypothetical prote
1128 4 14.3 46 2 D99138 hypothetical prote
1129 4 14.3 46 2 G99761 hypothetical prote
1130 4 14.3 48 2 AC1046 entericidin B prec
1131 4 14.3 49 2 D95054 hypothetical prote
1132 4 14.3 50 2 C60407 monocyte adherence
1133 4 14.3 50 2 D90403 hypothetical prote
1134 4 14.3 50 2 G82540 hypothetical prote
1135 4 14.3 51 2 B32040 dihydrolipoamide S
1136 4 14.3 51 2 F84017 hypothetical prote
1137 4 14.3 51 2 D97830 hypothetical prote
1138 4 14.3 51 2 D98076 hypothetical prote
1139 4 14.3 52 2 G71256 probable rubredoxi
1140 4 14.3 52 2 S21591 Ig heavy chain V r
1141 4 14.3 53 1 RGBPA2 regulatory protein
1142 4 14.3 53 2 S74979 hypothetical prote
1143 4 14.3 54 2 T11968 hypothetical prote
1144 4 14.3 54 2 C97093 hypothetical prote
1145 4 14.3 55 1 FECLCB ferredoxin 2[4Fe-4
1146 4 14.3 55 2 T11171 H+-transporting AT
1147 4 14.3 55 2 T11120 H+-transporting AT
1148 4 14.3 55 2 A24633 M protein - Strept
1149 4 14.3 55 2 E90616 ATP synthase F0 ch
1150 4 14.3 55 2 E81417 small hydrophobic
1151 4 14.3 55 2 JG6188 csGAB protein - Ba
1152 4 14.3 55 2 A37238 autoimmune epitope
1153 4 14.3 55 2 H97965 degenerate transpo
1154 4 14.3 56 2 S45028 sox2 protein - Her
1155 4 14.3 56 2 A95955 hypothetical expor
1156 4 14.3 57 2 B43603 antigen 85-C - Myc
1157 4 14.3 57 2 F81889 hypothetical prote
1158 4 14.3 57 2 S26205 metallothionein -
1159 4 14.3 57 2 I34529 major histocompati
1160 4 14.3 57 2 D82854 hypothetical prote
1161 4 14.3 57 2 E97063 hypothetical prote
1162 4 14.3 57 2 C97813 conserved hypothet
1163 4 14.3 57 2 AG3105 Ig heavy chain V r
1164 4 14.3 58 2 S31683 protein kinase Mpk
1165 4 14.3 59 2 S30504 ADP-ribosylation f
1166 4 14.3 59 2 A28080 retinoic acid rece
1167 4 14.3 59 2 I51272 hypothetical prote
1168 4 14.3 59 2 H90943 plxv protein - Pse
1169 4 14.3 59 2 T12018 hypothetical prote
1170 4 14.3 59 2 C64942 hypothetical prote
1171 4 14.3 59 2 T37158 hypothetical prote
1172 4 14.3 59 2 D85792 conserved hypothet
1173 4 14.3 59 2 AH0725 hypothetical prote
1174 4 14.3 60 1 QQP5HC hypothetical prote
1175 4 14.3 60 1 QQEBHC phosphoenolpyruvat
1176 4 14.3 60 2 T05702 hypothetical prote
1177 4 14.3 60 2 S77978 homeotic protein -
1178 4 14.3 60 2 S13012 homeotic protein -
1179 4 14.3 60 2 S13013 conserved protein,
1180 4 14.3 60 2 A10170 conserved hypothet
1181 4 14.3 60 2 F98317 hypothetical prote
1182 4 14.3 61 2 S13992 hypothetical prote
1183 4 14.3 61 2 S8741 protein YPK170w-a
1184 4 14.3 62 2 H82625 hypothetical prote
1185 4 14.3 63 1 H70394 ribosomal protein
1186 4 14.3 63 1 YVBP6A lysis protein - ph
1187 4 14.3 63 1 B46324 lysis protein - ph
1188 4 14.3 63 2 E99333 4-oxalocrotonate t
1189 4 14.3 63 2 G82268 hypothetical prote
1190 4 14.3 63 2 H70930 probable lppt prot
1191 4 14.3 63 2 S29654 juvenile-hormone e
1192 4 14.3 63 2 B58799 cannabinoid recept
1193 4 14.3 63 2 T30614 hypothetical prote
1194 4 14.3 64 2 S20895 homeotic protein c
1195 4 14.3 64 2 G87525 transcription regu
1196 4 14.3 64 2 S33489 hypothetical 6.7K
1197 4 14.3 64 2 E72361 hypothetical prote

1198 4 14.3 64 2 H75270 transcription regu
1199 4 14.3 64 2 S34660 probable cold-regu
1200 4 14.3 64 2 D97727 hypothetical prote

ALIGNMENTS

RESULT 1

T10442
selenoprotein P precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T10442
R:Steinert, P.; Ahrens, M.; Gross, G.; Flohe, L.
BioFactors 6, 311-319, 1997
A:Title: cDNA and deduced polypeptide sequence of a mouse selenoprotein P.
A:Reference number: Z17017; MUID:97434516
A:Accession: T10442
A:Status: Preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-380 <STE>
A:Cross-references: EMBL:X99807; NID:g1495998; PIDN:CAA68140.1; PID:g1495999
C:Genetics:
A:Gene: selp
C:Superfamily: selenoprotein P
C:Keywords: extracellular protein; selenocysteine

Query Match 39.3%; Score 11; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ENLPSCSXOGL 22

DB 322 ENLPSCSXOGL 333

RESULT 2

A47327
selenoprotein P precursor [validated] - human
C:Species: Homo sapiens (man)
C>Date: 21-Sep-1993 #sequence_revision 01-Dec-1995 #text_change 15-Sep-2000
C:Accession: A47327; S42752
R:Hill, K.E.; Lloyd, R.S.; Burk, R.F.
Proc. Natl. Acad. Sci. U.S.A. 90, 537-541, 1993
A:Title: Conserved nucleotide sequences in the open reading frame and 3' untranslated
A:Reference number: A47327; MUID:93133823
A:Accession: A47327
A:Molecule type: mRNA
A:Residues: 1-381 <HIL>
A:Cross-references: GB:Z11793; NID:g36425; PIDN:CAA77836.1; PID:g2654365
A:Experimental source: heart and liver
A:Note: In Genbank entry HSESLPM, release 117.0, PIDN:CAA77836.1, the selenocysteine
R:Akesson, B.; Bellw, T.; Burk, R.F.
Biochim. Biophys. Acta 1204, 243-249, 1994
A:Title: Purification of selenoprotein P from human plasma.
A:Reference number: S42752; MUID:94191007
A:Accession: S42752
A:Molecule type: protein
A:Residues: 20-27,'X',29-33 <AKE>
A:Note: mature forms of 55K and 61K were detected in plasma; the protein was shown to

C:Genetics: A:Gene: GDB:SEPP1; SLNP

A:Cross-references: GDB:138278; OMIM:601484

A:Map position: 5q31-5q31

C:Function: A:Description: may act as a free-radical scavenger

C:Superfamily: selenoprotein P

C:Keywords: extracellular protein; glycoprotein; heparin binding; liver; plasma; sele
F:1-15/Domain: signal sequence #status predicted <SIG>
F:20-381/Product: selenoprotein P #status experimental <MAT>
F:46,83,119,128,338/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

F:59,300,318,330,345,352,367,369,376,378/Modified site: selenocysteine #status predicted

Query Match      39.3%; Score 11; DB 1; Length 381;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 OCKENLPSLCS 18
    |||||
Db 319 OCKENLPSLCS 329

RESULT 3
OMRSP
selenoprotein P precursor [validated] - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 15-Sep-2000
C:Accession: A40380; B40380; S68322
R: Hill, K.E.; Lloyd, R.S.; Yang, J.G.; Read, R.; Burk, R.F.
J. Biol. Chem. 266, 10050-10053, 1991
A:Title: The cDNA for rat selenoprotein P contains 10 TGA codons in the open reading frame
A:Reference number: A40380; MUID:91244760
A:Accession: A40380
A:Molecule type: mRNA
A:Residues: 1-385 <H1>
A:Cross-references: GB:M63574; NID:g206893; PIDN:AAA42129.1; PID:g206894
A:Accession: B40380
A:Molecule type: protein
A:Residues: 20-41; 267-287; 316-327 <H12>
R: Chittum, H.S.; Himeno, S.; Hill, K.E.; Burk, R.F.
Arch. Biochem. Biophys. 325, 124-128, 1996
A:Title: Multiple forms of selenoprotein P in rat plasma.
A:Reference number: S68322; MUID:96140605
A:Accession: S68322
A:Molecule type: protein
A:Residues: 20-27 <CHI>
C:Superfamily: selenoprotein P
C:Keywords: extracellular protein; glycoprotein; heparin binding; liver; plasma; selenocysteine
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-385/Product: selenoprotein P #status experimental <MAT>
F:59,264,335,357,371,373,380,382/Modified site: selenocysteine #status predicted
F:83,174,188,370,375/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:282,323/Modified site: selenocysteine #status experimental

Query Match      28.6%; Score 8; DB 1; Length 385;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 ENLPSLCS 18
    |||||
Db 327 ENLPSLCS 334

RESULT 4
F84151
Response regulator aspartate phosphatase BH4014 [imported] - Bacillus halodurans (strain
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: F84151
R: Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: F84151
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-370 <STO>
A:Cross-references: GB:AP001520; GB:BA000004; NID:g10176401; PIDN:BA07733.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH4014

```

```

Query Match      25.0%; Score 7; DB 2; Length 370;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KENLPSL 16
    |||||
Db 357 KENLPSL 363

```

```

RESULT 5
AC0673
conserved hypothetical protein STY1496 [imported] - Salmonella enterica subsp. enteri
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: AC0673
R: Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Connerly, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Ouail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
A:Reference number: AB0502; PMID:11677608
A:Accession: AC0673
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-71 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD01756.1; PID:g16502604; GSPDB:GN00176
C:Genetics:
A:Gene: STY1496

```

```

Query Match      21.4%; Score 6; DB 2; Length 71;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 QGLRAE 25
    |||||
Db 22 QGLRAE 27

```

```

RESULT 6
E90889
hypothetical protein ECS2085 [imported] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: E90889
R: Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: E90889
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-92 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA035508.1; PID:g13361551; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECS2085

```

```

Query Match      21.4%; Score 6; DB 2; Length 92;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 QGLRAE 25
    |||||
Db 43 QGLRAE 48

```

```

RESULT 7

```

D85728
hypothetical protein 22229 [imported] - Escherichia coli (strain O157:H7, substrain EDL99)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: D85728
R:Perina, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouisis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85728
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-92 <STO>
A:Cross-references: GB:AE005174; NID:g12515200; PIDN:AAG56288.1; GSPDB:GN00145; UWGP:222
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: 72229

Query Match 21.4%; Score 6; DB 2; Length 92;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
|||||
Db 43 QGLRAE 48

RESULT 8
hypothetical protein Sma1100 [imported] - Sinorhizobium meliloti (strain 1021) magaplas
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: F95336
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: F95336
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-153 <KUR>
A:Cross-references: GB:AE006469; PIDN:AAK65256.1; PID:g14523707; GSPDB:GN00165
A:Experimental source: strain 1021, megaplasmid pSymA
R:Galibert, F.; Finan, T.M.; Long, S.R.; Punler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Leilaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: Sma1100
A:Genome: plasmid

Query Match 21.4%; Score 6; DB 2; Length 153;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GLRAEE 26
|||||
Db 22 GLRAEE 27

RESULT 9
G83712
hypothetical protein BH0503 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: G83712
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: G83712
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-174 <STO>
A:Cross-references: GB:AP001508; GB:BA000004; NID:g10172890; PIDN:BA04222.1; GSPDB:G
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH0503
C:Superfamily: Deinococcus radiodurans hypothetical protein DR0763

Query Match 21.4%; Score 6; DB 2; Length 174;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KENLPS 15
|||||
Db 138 KENLPS 143

RESULT 10
G69549
acetyl-CoA decarboxylase/synthase, subunit epsilon (cdhB-2) homolog - Archaeoglobus f
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
C:Accession: G69549
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E
Glodek, A.; Zhou, L.; Overbeek, R.; Gayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes,
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch
A:Reference number: A69250; MUID:98049343
A:Accession: G69549
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-175 <KLE>
A:Cross-references: GB:AE001110; GB:AE000782; NID:g2689433; PIDN:AAB91265.1; PID:g265
C:Superfamily: carbon-monoxide dehydrogenase beta chain MTH1709

Query Match 21.4%; Score 6; DB 2; Length 175;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
|||||
Db 68 TGSALT 73

RESULT 11
S76419
probable ferredoxin [2Fe-2S] - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C:Accession: S76419
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys
s.
A:Reference number: S74322; MUID:97061201
A:Accession: S76419
A:Status: preliminary
A:Molecule type: DNA

A;Residues: 1-186 <KAN>
A;Cross-references: EMBL:D90915; GB:AB001339; NID:g1653604; PIDN:BAAL8548.1; PID:d101928
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Superfamily: ferredoxin [2Fe-2S] homology
C;Keywords: 2Fe-2S; iron-sulfur protein; metalloprotein
F:48-109/Domain: ferredoxin [2Fe-2S] homology <FER>
F:63,69,72,108/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 21.4%; Score 6; DB 2; Length 186;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 RAENI 28
|||||
DB 140 RAENI 145

RESULT 12
D83125
probable proteinase PA4171 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: D83125
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.: Lory, S.; Olson, M.V.
A;Residues: 1-187 <STO>
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A;Reference number: A82950; MUID:20437337
A;Accession: D83125
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-187 <STO>
A;Cross-references: GB:AE004833; GB:AE004091; NID:g9950370; PIDN:AG07558.1; GSPDB:GN001
A;Experimental source: strain PAO1
C;Genetics:
C;Superfamily: Archaeoglobus intracellular proteinase I

Query Match 21.4%; Score 6; DB 2; Length 187;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
|||||
DB 68 QGLRAE 73

RESULT 13
D90040
hypothetical protein SA2183 [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: D90040
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogo
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: D90040
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-191 <KUR>
A;Cross-references: GB:BA000018; PID:g13702344; PIDN:BA843485.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:
A;Gene: SA2183

Query Match 21.4%; Score 6; DB 2; Length 191;

C:Accession: F81386
 F:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barre
 Nature 403, 665-668, 2000
 A:Title: The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hyp
 A:Reference number: A81250; MUID:20150912
 A:Accession: F81386
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-210 <PAR>
 A:Cross-references: GB:AL111168; NID:g6967817; PIDN:CAB74260.1; PID:g696789
 A:Experimental source: serotype O2, strain NCTC 11168
 C:Genetics:
 A:Gene: Cj0424

Query Match 21.4%; Score 6; DB 2; Length 210;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 LRAEEN 27
 |||||
 Db 67 LRAEEN 72

RESULT 17
 S39522
 hypothetical protein X - red alga (*Cyanidium caldarium*) chloroplast (fragment)
 C:Species: Chloroplast *Cyanidium caldarium*
 C:Date: 07-Oct-1994 #sequence_revision 14-Nov-1997 #text_change 29-Sep-1999
 C:Accession: S39522
 R:Kostorzewa, M.; Zetsche, K.
 Plant Mol. Biol. 23, 67-76, 1993
 A:Title: Organization of plastid-encoded ATPase genes and flanking regions including hom
 A:Reference number: S39512; MUID:94033298
 A:Accession: S39522
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-221 <KOS>
 A:Cross-references: EMBL:X67814; NID:g429169; PIDN:CAA48016.1; PID:g429180
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1992
 A:Note: the source is designated as *Galdieria sulphuraria*
 C:Genetics:
 A:Genome: chloroplast
 C:Superfamily: conserved hypothetical protein b1683
 C:Keywords: chloroplast

Query Match 21.4%; Score 6; DB 2; Length 221;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGSAT 6
 |||||
 Db 47 TGSAT 52

RESULT 18
 G86387
 hypothetical protein F2823.16 - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
 C:Accession: G86387
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A:Reference number: A86141; MUID:21016719
 A:Accession: G86387
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-248 <STO>
 A:Cross-references: GB:AE005172; NID:g1079509; PIDN:AAG29220.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1

Query Match 21.4%; Score 6; DB 2; Length 248;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 GLRAEE 26
 |||||
 Db 165 GLRAEE 170

RESULT 19
 A97227
 polyferredoxin [imported] - *Clostridium acetobutylicum*
 C:Species: *Clostridium acetobutylicum*
 C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
 C:Accession: A97227
 R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L
 J. Bacteriol. 183, 4833-4838, 2001
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
 A:Reference number: A96900; MUID:21359325; PMID:21359325
 A:Accession: A97227
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-249 <KUP>
 A:Cross-references: GB:AE001437; PIDN:AAK80604.1; PID:g15025686; GSPDB:GN00168
 A:Experimental source: *Clostridium acetobutylicum* ATCC824
 C:Genetics:
 A:Gene: CAC2657

Query Match 21.4%; Score 6; DB 2; Length 249;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KENLPS 15
 |||||
 Db 70 KENLPS 75

RESULT 20
 S39737
 glucose 1-dehydrogenase homolog ywfd - *Bacillus subtilis*
 N:Alternate names: protein ipa-82d
 N:Contains: probable dehydrogenase (EC 1.1.1.-)
 C:Species: *Bacillus subtilis*
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
 C:Accession: S39737; E70055
 R:Glaser, P.; Kunst, F.; Arnaud, M.; Coudart, M.P.; Gonzales, W.; Hullo, M.F.; Ionesc
 , A.; Rapoport, G.; Danchin, A.
 Mol. Microbiol. 10, 371-384, 1993
 A:Title: *Bacillus subtilis* genome project: cloning and sequencing of the 97 kb region
 A:Reference number: S39655; MUID:95020537

A:Accession: S39737
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-255 <GLA>
 A:Cross-references: EMBL:X73124; NID:g413923; PIDN:CAA51638.1; PID:g414006
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1993
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferraril,
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal

iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.P.; Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y.; M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli, R.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, A.; Authors: Schleicher, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:98044033
A:Accession: E70055
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-255 <KUR>
A:Cross-references: GB:299123; GB:AL009126; NID:g2636240; PIDN:CAB15799.1; PID:g2636308
A:Experimental source: strain 168
C:Genetics:
A:Gene: ywifD
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
C:Keywords: NAD; oxidoreductase
F:8-185/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 21.4%; Score 6; DB 1; Length 255;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
| | | | |
Db 242 TGSALT 247

RESULT 21
H95954
C:Species: *Sinorhizobium meliloti*
Probable Sir2-like transcription silencer protein [imported] - *Sinorhizobium meliloti* (S)
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: H95954
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb pSymb megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: H95954
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-272 <KUR>
A:Cross-references: GB:AL591985; PIDN:CAC49304.1; PID:g15140790; GSPDB:GNO0167
A:Experimental source: strain 1021, megaplasmid pSymb
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
peia, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: Smb21328
A:Genome: plasmid

Query Match 21.4%; Score 6; DB 2; Length 272;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QGLRAE 25
| | | | |
Db 264 QGLRAE 269

RESULT 22

AG25332
hypothetical protein alr7546 [imported] - *Anabaena* sp. (strain PCC 7120) plasmid pCC7
C:Species: *Anabaena* sp.
A:Note: *Anabaena* sp. (strain PCC 7120) is a synonym of *Nostoc* sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C:Accession: AG2532
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing *Cyanobacterium*
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG2532
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-276 <KUR>
A:Cross-references: GB:AP003602; PIDN:BA877189.1; PID:g17134631; GSPDB:GNO0181
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr7546
A:Genome: plasmid

Query Match 21.4%; Score 6; DB 2; Length 276;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KENLPS 15
| | | | |
Db 160 KENLPS 165

RESULT 23
G24723
C:Species: *Corynebacterium glutamicum*
tryptophan synthase (EC 4.2.1.20) alpha chain - *Corynebacterium glutamicum*
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
C:Accession: G24723; C29458
R:Matsui, K.; Sano, K.; Ohtsubo, E.
Nucleic Acids Res. 14, 10113-10114, 1986
A:Title: Complete nucleotide and deduced amino acid sequences of the *Brevibacterium* 1
A:Reference number: A93606; MUID:87117512
A:Accession: G24723
A:Molecule type: DNA
A:Residues: 1-281 <MAT>
A:Cross-references: GB:X04960; NID:g39591; PIDN:CAA28628.1; PID:g39598
A:Note: the source is designated as *Brevibacterium lactofermentum*
R:Sano, K.; Matsui, K.
Gene 53, 191-200, 1987
A:Title: Structure and function of the trp operon control regions of *Brevibacterium* 1
A:Reference number: A91575; MUID:87277409
A:Accession: C29458
A:Molecule type: DNA
A:Residues: 161-281 <SAN>
A:Cross-references: GB:M16664; NID:g144097; PIDN:AAA83990.1; PID:g144100
A:Note: the source is designated as *Brevibacterium lactofermentum*
C:Genetics:
A:Gene: trpA
C:Superfamily: tryptophan synthase alpha chain; tryptophan synthase alpha chain homol
C:Keywords: carbon-oxygen lyase; hydro-lyase
F:19-245/Domain: tryptophan synthase alpha chain homology <TRPA>
F:50/Active site: Glu #status predicted

Query Match 21.4%; Score 6; DB 2; Length 281;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
| | | | |
Db 232 TGSALT 237

RESULT 24

S37636
 hypothetical protein 3 (atpA 3' region) - red alga (Antithamnion sp.) chloroplast (fragment)
 C:Species: chloroplast Antithamnion sp.
 C:Date: 07-Oct-1994 #sequence_revision 30-Jan-1998 #text_change 29-Sep-1999
 C:Accession: S37636
 R:Kostrzewa, M.; Zetsche, K.
 J. Mol. Biol. 227, 961-970, 1992
 A:Title: Large ATP synthase operon of the red alga Antithamnion sp. resembles the corresponding operon of the green alga *Chlorella*
 A:Reference number: S26957; MUID:93021132
 A:Accession: S37636
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-297 <KOS>
 A:Cross-references: EMBL:X63382; NID:g14170; PIDN:CAA44986.1; PID:g14179
 A:Experimental source: strain LB 95.79
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1991
 C:Genetics:
 A:Genome: chloroplast
 C:Superfamily: conserved hypothetical protein b1683
 C:Keywords: chloroplast

Query Match 21.4%; Score 6; DB 2; Length 297;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSAT 6
 |||||
 Db 123 TGSAT 128

RESULT 25
 JS0681
 hypothetical protein 4 (glda 5' region) - Bacillus stearothermophilus
 C:Species: Bacillus stearothermophilus
 C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 15-Oct-1999
 C:Accession: JS0681
 R:Mailinder, P.R.; Pritchard, A.; Moir, A.
 Gene 110, 9-16, 1992
 A:Title: Cloning and characterization of a gene from Bacillus stearothermophilus var. n. encoding a protein with homology to the protein encoded by the *glda* gene of *Bacillus stearothermophilus*
 A:Reference number: JQ1474; MUID:92184120
 A:Accession: JS0681
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-301 <MAL>
 A:Cross-references: GB:M65289; NID:g142976; PIDN:AAA22479.1; PID:g142980
 A:Experimental source: strain var. non-diastaticus
 C:Genetics:
 A:Start codon: GTG

Query Match 21.4%; Score 6; DB 2; Length 301;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GLRAEE 26
 |||||
 Db 218 GLRAEE 223

RESULT 26
 S61033
 hypothetical protein YDL179w - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein D1408
 C:Species: Saccharomyces cerevisiae
 C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 20-Jun-2000
 C:Accession: S61033; S67731
 R:Pohl, T.M.
 submitted to the EMBL Data Library, November 1995
 A:Reference number: S61010
 A:Accession: S61033
 A:Molecule type: DNA
 A:Residues: 1-304 <POH>

A:Cross-references: EMBL:Z67750; NID:g1061256; PIDN:CAA91566.1; PID:g1061259
 R:Pohl, T.M.
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S67708
 A:Accession: S67731
 A:Molecule type: DNA
 A:Residues: 1-304 <POW>
 A:Cross-references: EMBL:Z74227; NID:g1431288; PIDN:CAA98753.1; PID:g1431289; MIPS:YD
 A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: SGD:PCL9
 A:Cross-references: SGD:S0002338; MIPS:YDL179w
 A:Map position: 4L

Query Match 21.4%; Score 6; DB 2; Length 304;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 RAEENI 28
 |||||
 Db 260 RAEENI 265

RESULT 27
 B69553
 methanol dehydrogenase regulatory protein (moxR) homolog - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: B69553
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod
 .; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Uitterlinden, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes,
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch
 A:Reference number: A69250; MUID:98049343
 A:Accession: B69553
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-324 <KLE>
 A:Cross-references: GB:AE001108; GB:AE000782; NID:g2689431; PIDN:AAB91247.1; PID:g265
 C:Superfamily: methanol dehydrogenase regulatory protein

Query Match 21.4%; Score 6; DB 1; Length 324;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GLRAEE 26
 |||||
 Db 300 GLRAEE 305

RESULT 28
 T20729
 hypothetical protein F10G8.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
 C:Accession: T20729
 R:Basham, V.
 submitted to the EMBL Data Library, September 1996
 A:Reference number: Z19315
 A:Accession: T20729
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-352 <WIL>
 A:Cross-references: EMBL:Z80216; PIDN:CAB02287.1; GSPDB:GN00019; CESP:F10G8.1
 A:Experimental source: clone F10G8
 C:Genetics:
 A:Gene: CESP:F10G8.1
 A:Map position: 1

A: Introns: 5/1; 40/3; 192/3; 257/2; 309/3
C: Superfamily: protein-tyrosine-phosphatase, nonreceptor type 12; protein-tyrosine-phosphatase

Query Match 21.4%; Score 6; DB 2; Length 352;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 OQLRAE 25
|||||
DB 54 OQLRAE 59

RESULT 29
150712
TOP AP - chicken
C: Species: Gallus gallus (chicken)
C: Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C: Accession: I50712
R: Savitt, J.M.; Trisler, D.; Hilt, D.C.
Neuron 14, 253-261, 1995
A: Title: Molecular cloning of TOPAP: a topographically graded protein in the developing
A: Reference number: I50712; MUID: 95161057
A: Accession: I50712
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: mRNA
A: Residues: 1-359 <SAV>
A: Cross-references: EMBL:U17000; NID:9642485; PIDN:AAA68637.1; PID:9642486

Query Match 21.4%; Score 6; DB 2; Length 359;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 OCKENL 13
|||||
DB 318 OCKENL 323

RESULT 30
H71324
probable uracil phosphoribosyltransferase - syphilis spirochete
C: Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C: Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
C: Accession: H71324
R: Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwirdson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDowell, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A: Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A: Reference number: A71250; MUID: 98332770
A: Accession: H71324
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-360 <COL>
A: Cross-references: GB:AF001221; GB:AF0000520; NID:93322720; PIDN:AAC65434.1; PID:9332273
A: Experimental source: strain Nichols
C: Genetics:
A: Gene: TP0448

Query Match 21.4%; Score 6; DB 2; Length 360;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 RAEENI 28
|||||
DB 56 RAEENI 61

RESULT 31
A45999
tyrosine--tRNA ligase (EC 6.1.1.1) [validated] - yeast (Saccharomyces cerevisiae)

N: Alternate names: protein G7522; protein YGR185c; tyrosyl-tRNA synthetase
C: Species: Saccharomyces cerevisiae
C: Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 20-Jun-2000
C: Accession: A45999; S64503
R: Chow, C.M.; RajBhandary, U.L.
J. Biol. Chem. 268, 12855-12863, 1993
A: Title: Saccharomyces cerevisiae cytoplasmic tyrosyl-tRNA synthetase gene. Isolation
A: Reference number: A45999; MUID: 93286133
A: Accession: A45999
A: Molecule type: DNA
A: Residues: 1-394 <CHO>
A: Cross-references: GB:L12221; NID:gl73119; PIDN:AAB59329.1; PID:gl73120
R: Arroyo, J.; Garcia-Gonzalez, M.; Garcia-Saez, M.I.; Sanchez-Perez, M.; Nombela, C.
submitted to the Protein Sequence Database, May 1996
A: Reference number: S64499
A: Accession: S64503
A: Molecule type: DNA
A: Residues: 1-394 <ARR>
A: Cross-references: EMBL:Z72970; NID:gl1323327; PIDN:CAA97211.1; PID:gl1323328; GSPDB:G
A: Experimental source: strain S288C
C: Genetics:
A: Gene: SGD-TYS1; MGW104; MIPS:YGR185c
A: Cross-references: SGD:S0003417; MIPS:YGR185c
A: Map position: 7R
C: Function:
A: Description: EC 6.1.1.1 [validated, MUID:93286133]; catalyzes the ATP-dependent for
A: Pathway: protein biosynthesis
C: Superfamily: yeast tyrosine--tRNA ligase
C: Keywords: aminocyl-tRNA synthetase; ATP; ligase; protein biosynthesis

Query Match 21.4%; Score 6; DB 2; Length 394;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ENLPSL 16
|||||
DB 201 ENLPSL 206

RESULT 32
S72814
hypothetical protein B1620_C2_213 - Mycobacterium leprae
C: Species: Mycobacterium leprae
C: Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 23-Mar-2001
C: Accession: S72814
R: Smith, D.R.; Robison, K.
submitted to the EMBL Data Library, November 1993
A: Description: Mycobacterium leprae cosmid B1620.
A: Reference number: S72584
A: Accession: S72814
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-394 <SMI>
A: Cross-references: EMBL:U00015; NID:9466931; PIDN:AAC43223.1; PID:9466935

Query Match 21.4%; Score 6; DB 2; Length 394;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ENLPSL 16
|||||
DB 274 ENLPSL 279

RESULT 33
A46699
translation initiation factor eIF-5 - yeast (Saccharomyces cerevisiae)
N: Alternate names: protein YP3085.05; protein YPR041w
C: Species: Saccharomyces cerevisiae
C: Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 06-Feb-1998

C:Accession: A46699; B46699; S61062
R:Chakravarti, D.; Maitra, U.
J. Biol. Chem. 268, 10524-10533, 1993
A:Title: Eukaryotic translation initiation factor 5 from *Saccharomyces cerevisiae*. Cloned
A:Reference number: A46699; MUID:93252941
A:Accession: A46699
A:Molecule type: DNA
A:Residues: 1-405 <CHA>
A:Cross-references: EMBL:L10840
A:Note: sequence extracted from NCBI backbone (NCBIN:131322, NCBIPI:131325)
A:Accession: B46699
A:Molecule type: protein
A:Residues: 19-23,'X',25,'X',27-29;326,'X',328-331,'X',333-337,'X',339-340;363-368,'X',3
R:Badcock, K.; Churcher, C.M.
submitted to the EMBL Data Library, November 1995
A:Reference number: S61058
A:Accession: S61062
A:Molecule type: DNA
A:Residues: 1-405 <BAD>
A:Cross-references: EMBL:Z68111; NID:g1072402; PID:g1072407; MIPS:YPR041w
C:Genetics:
A:Gene: SGD:TIF5
A:Cross-references: SGD:S0006245; MIPS:YPR041w
A:Map position: 16R
C:Function:
A:Description: protein biosynthesis
C:Keywords: alternative initiators; protein biosynthesis
F:1-405/Product: translation initiation factor eIF-5 long form #status predicted <WAT>
F:18-405/Product: translation initiation factor eIF-5 short form #status predicted <MAT>

Query Match 21.4%; Score 6; DB 2; Length 405;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KENLPS 15
|||||
Db 263 KENLPS 268

RESULT 34
T23954
Hypothetical protein R06B9.1 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C:Accession: T23954
R:Baynes, C.
submitted to the EMBL Data Library, December 1996
A:Reference number: Z19823
A:Accession: T23954
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-427 <WIL>
A:Cross-references: EMBL:Z83237; PIDN:CAB05786.1; GSPDB:GN00020; CESP:R06B9.1
A:Experimental source: clone R06B9
C:Genetics:
A:Gene: CESP:R06B9.1
A:Map position: 2
A:Introns: 62/2; 97/3; 156/1; 380/1
C:Superfamily: *Caenorhabditis elegans* hypothetical protein R06B9.1

Query Match 21.4%; Score 6; DB 2; Length 427;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 LRAEN 27
|||||
Db 397 LRAEN 402

RESULT 35
G83350

hypothetical protein PA2363 [imported] - *Pseudomonas aeruginosa* (strain PA01)
C:Species: *Pseudomonas aeruginosa*
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: G83350
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337
A:Accession: G83350
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-443 <STO>
A:Cross-references: GB:AE004662; GB:AE004091; NID:g9948395; PIDN:AAG05751.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA2363

Query Match 21.4%; Score 6; DB 2; Length 443;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 OGLRAE 25
|||||
Db 24 OGLRAE 29

RESULT 36
S36412
H+-transporting ATP synthase (EC 3.6.1.34) beta chain - red alga (*Cyanidium caldarium*
C:Species: *Cyanidium caldarium*
C:Date: 10-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 19-Jan-2001
C:Accession: S36412
R:Kostrzewa, M.; Zetsche, K.
submitted to the EMBL Data Library, June 1992
A:Reference number: S36410
A:Accession: S36412
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-476 <KOS>
A:Superfamily: H+-transporting ATP synthase alpha chain; H+-transporting ATP synthase
C:Keywords: ATP biosynthesis; chloroplast; hydrolase; nucleotide binding; P-loop
F:155-162/Region: nucleotide-binding motif A (P-loop)
F:181-357/Domain: H+-transporting ATP synthase alpha chain homology <ATP>

Query Match 21.4%; Score 6; DB 2; Length 476;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSAIT 6
|||||
Db 77 TGSAIT 82

RESULT 37
AE2117
hypothetical protein alr2492 [imported] - *Anabaena* sp. (strain PCC 7120)
C:Species: *Anabaena* sp.
A:Note: *Anabaena* sp. (strain PCC 7120) is a synonym of *Nostoc* sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C:Accession: AE2117
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing *Cyanobacterium*
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AE2117
A>Status: preliminary
A:Molecule type: DNA

Query Match 21.4%; Score 6; DB 2; Length 485;
C;Superfamily: conserved hypothetical protein b1683

Best Local Similarity 100.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
|||||
DB 311 TGSALT 316

RESULT 42

T06907

probable ABC transport protein - Cyanophora paradoxa cyanelle

C:Species: cyanelle Cyanophora paradoxa

C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 29-Sep-1999

C:Accession: T06907

R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohnert, H.J.; Bryant, D.A.

submitted to the EMBL Data Library, July 1995

A:Description: Nucleotide sequence of the cyanelle genome from Cyanophora paradoxa.

A:Reference number: Z15840

A:Accession: T06907

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-486 <ST>

A:Cross-references: EMBL:U30821; NID:g1016083; PIDN:AA81250.1; PID:g1016163

A:Experimental source: strain Pringsheim LB555

C:Genetics:

A:Gene: ycf24

A:Genome: cyanelle

C:Superfamily: conserved hypothetical protein b1683

C:Keywords: cyanelle

Query Match 21.4%; Score 6; DB 2; Length 486;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
|||||
DB 312 TGSALT 317

RESULT 43

S07062

glutamate receptor precursor - chicken

N:Alternate names: kainate-binding protein

C:Species: Gallus gallus (chicken)

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: S07062; S27105

R:Gregor, P.; Mano, I.; Maoz, I.; McKeown, M.; Teichberg, V.I.

Nature 342, 689-692, 1989

A:Title: Molecular structure of the chick cerebellar kainate-binding subunit of a putative

A:Reference number: S07062; MUID:90081841

A:Accession: S07062

A:Molecule type: DNA

A:Residues: 1-487 <GRE>

A:Cross-references: EMBL:X17700; NID:g62786; PIDN:CRA35693.1; PID:g62787

A:Accession: S27105

A:Molecule type: protein

A:Residues: 24-27, 'X', 29-30, 'X', 32-37, 61, 'X', 63-70, 'X', 72-74, 'XX', 77-81, 'X', 83-90, 'X', 92

C:Superfamily: kainate-binding protein; glutamate receptor homology

C:Keywords: glycoprotein; ion channel; neurotransmitter receptor; phosphoprotein; trans

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-487/Product: glutamate receptor #status experimental <MAT>

F:51-454/Domain: glutamate receptor homology <GRP>

F:172-190/Domain: transmembrane #status predicted <TM1>

F:207-225/Domain: transmembrane #status predicted <TM2>

F:236-260/Domain: transmembrane #status predicted <TM3>

F:419-440/Domain: transmembrane #status predicted <TM4>

F:104, 444/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:380, 408/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 21.4%; Score 6; DB 1; Length 487;
Best Local Similarity 100.0%; Pred. No. 77;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ENLPSL 16
|||||
DB 44 ENLPSL 49

RESULT 44

S73161

hypothetical protein 24 - red alga (Porphyra purpurea) chloroplast

C:Species: chloroplast Porphyra purpurea

C>Date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 29-Sep-1999

C:Accession: S73161

R:Reith, M.; Munholland, J.

Plant Mol. Biol. Rep. 13, 333-335, 1995

A:Title: Complete nucleotide sequence of the Porphyra purpurea chloroplast genome.

A:Reference number: S73108

A:Accession: S73161

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-487 <REI>

A:Cross-references: EMBL:U38804; NID:g1276652; PIDN:AAC08126.1; PID:g1276706

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995

C:Genetics:

A:Gene: ycf24

A:Genome: chloroplast

C:Superfamily: conserved hypothetical protein b1683

C:Keywords: chloroplast

Query Match 21.4%; Score 6; DB 2; Length 487;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
|||||
DB 313 TGSALT 318

RESULT 45

C87480

conserved hypothetical protein CCl864 [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: C87480

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,

B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: C87480

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-489 <STO>

A:Cross-references: GB:AE005673; NID:gl3423307; PIDN:AAK23839.1; GSPDB:GN00148

C:Genetics:

A:Gene: CCl864

Query Match 21.4%; Score 6; DB 2; Length 489;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
|||||
DB 315 TGSALT 320

RESULT 46

AF2800

ABC transporter subunit sufB [imported] - Agrobacterium tumefaciens (strain C58, Dupo

C:Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C;Accession: AF2800
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, H.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Rarp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
A;Accession: AF2800
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-489 <KUR>
A;Cross-references: GB:AE008688; PIDN:RAL42820.1; PID:g17740267; GSPDB:GN00186
A;Experimental source: Strain C58 (Dupont)
C;Genetics:
A;Gene: sufB
A;Map position: circular chromosome

Query Match 21.4%; Score 6; DB 2; Length 489;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
|||||
Db 315 TGSALT 320

RESULT 47
AB0703
conserved hypothetical protein STY1753 [imported] - Salmonella enterica subsp. enterica
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C;Accession: AB0703
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; PMID:11677608
A;Accession: AB0703
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-495 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD01995.1; PID:g16502833; GSPDB:GN00176
C;Genetics:
A;Gene: STY1753

Query Match 21.4%; Score 6; DB 2; Length 495;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
|||||
Db 321 TGSALT 326

RESULT 48
B83364
pyoverdine biosynthesis protein PvcC PA2256 [imported] - Pseudomonas aeruginosa (strain
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: B83364
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A;Reference number: AB2950; MUID:20437337
A;Accession: B83364
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-500 <STO>
A;Cross-references: GB:AE004651; GB:AE004091; NID:g9948277; PIDN:AAG05644.1; GSPDB:GN
C;Experimental source: strain PA01
C;Genetics:
A;Gene: pvcC; PA2256
C;Superfamily: Escherichia coli 4-hydroxyphenylacetate 3-monooxygenase large chain

Query Match 21.4%; Score 6; DB 2; Length 500;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
|||||
Db 189 TGSALT 194

RESULT 49
AD0293
conserved hypothetical protein YPO2403 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C;Accession: AD0293
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrel
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AD0293
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-503 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC91208.1; PID:g15980398; GSPDB:GN00175
C;Genetics:
A;Gene: YPO2403

Query Match 21.4%; Score 6; DB 2; Length 503;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGSALT 6
|||||
Db 329 TGSALT 334

RESULT 50
AD3382
ABC transporter-associated protein BMEI1042 [imported] - Brucella melitensis (strain
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C;Accession: AD3382
R;DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov
; Mazur, M.; Goltzman, E.; Seikov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melit
A;Reference number: AD3252; PMID:11756688
A;Accession: AD3382
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-507 <KUR>
A;Cross-references: GB:AE008917; PIDN:AAL52223.1; PID:g17983006; GSPDB:GN00190
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI1042
A;Map position: I

Query Match 21.4%; Score 6; DB 2; Length 507;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGSAIT 6
|
|
|
|
|
|
Db 333 TGSAIT 338

Search completed: August 22, 2002, 11:18:33
Job time: 100 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 07:49:22 ; Search time 13.59 Seconds
(without alignments)
82.624 Million cell updates/sec

Title: US-09-856-199-1

Perfect score: 164
Sequence: 1 KRGINQLLKLPTDSELAPRSXCHCRHL 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	161	98.2	381	1 SELL_HUMAN	P49908 homo sapien
2	122	74.4	380	1 SELL_MOUSE	P70274 mus musculus
3	116	70.7	402	1 SELL_BOVIN	P49907 bos taurus
4	107	65.2	385	1 SELL_RAT	P25236 rattus norv
5	53	32.3	324	1 Y734_CHLPN	O927h1 chlamydia p
6	52.5	32.0	335	1 LEP4_MVXXA	O30387 myxococcus
7	51.5	31.4	248	1 TPIS_RAT	P48500 rattus norv
8	51.5	31.4	521	1 VENV_DHVII	P27427 dhori virus
9	51	31.1	534	1 CNRD_CAEEL	P41830 caenorhabdi
10	50	30.5	588	1 ASH1_YEAST	P34233 saccharomyc
11	49.5	30.2	248	1 LEP4_BACSU	P15378 bacillus su
12	49	29.9	754	1 PURL_MYCTU	P54876 mycobacteri
13	48	29.3	290	1 LEP4_AERHY	P45794 aeromonas h
14	48	29.3	1297	1 IRR_HUMAN	P14616 homo sapien
15	47	28.7	83	1 DIS1_BITAR	P17497 bitis ariet
16	47	28.7	659	1 RNB_HAEIN	P44440 haemophilus
17	47	28.7	1286	1 KDV5_YEAST	P28273 saccharomyc
18	47	28.7	1339	1 ADO_BOVIN	P48034 bos taurus
19	46	28.0	119	1 TAT_SIVAG	P27982 simian immu
20	46	28.0	495	1 MRP2_DROME	Q24400 drosophila
21	46	28.0	507	1 ALSA_HUMAN	P15289 homo sapien
22	46	28.0	553	1 SYE_METTH	O26157 methanobact
23	46	28.0	747	1 EZH1_HUMAN	Q92800 homo sapien
24	46	28.0	747	1 EZH1_MOUSE	P70351 mus musculus
25	46	28.0	824	1 AD08_HUMAN	P78325 homo sapien
26	45.5	27.7	275	1 LEP4_BURPS	Q92f70 burkholderi
27	45.5	27.7	1801	1 LMB2_RAT	Q9da01 mus musculus
28	45	27.4	134	1 EPPI_MOUSE	P15021 drosophila
29	45	27.4	172	1 STEL_DROME	O09599 schizosacch
30	45	27.4	632	1 Y28_SCHPO	Q05910 mus musculus
31	45	27.4	826	1 AD08_MOUSE	O15230 homo sapien
32	45	27.4	3695	1 LMA5_HUMAN	P28797 cavia porce
33	44.5	27.1	591	1 GRN_CAVPO	

ALIGNMENTS

RESULT 1

34	44	26.8	63	1	MT2_COLLI	P15787 columba liv
35	44	26.8	84	1	AN11_HUMAN	Q9nyg5 homo sapien
36	44	26.8	84	1	AN11_MOUSE	Q9cpx9 mus musculus
37	44	26.8	112	1	COL_CANFA	P19090 canis famil
38	44	26.8	112	1	YN97_YEAST	P53977 saccharomyc
39	44	26.8	117	1	TAT_SIVS4	P12513 simian immu
40	44	26.8	128	1	TAT_SIVSP	P19507 simian immu
41	44	26.8	291	1	LEP4_AERSA	O68964 aeromonas s
42	44	26.8	388	1	VENV_MCV2	P25392 molluscum c
43	44	26.8	1338	1	ADO_HUMAN	Q06278 homo sapien
44	44	26.8	1912	1	CHD4_HUMAN	Q14839 homo sapien
45	44	26.8	2167	1	BEM2_YEAST	P39960 saccharomyc
46	43.5	26.5	183	1	US06_HOMVA	P14334 human cytom
47	43.5	26.5	269	1	LEP4_SYNY3	P72400 synecocyst
48	43.5	26.5	593	1	VE1_HPV48	Q80922 human papil
49	43.5	26.5	726	1	AD20_HUMAN	Q43506 homo sapien
50	43.5	26.5	1005	1	EP48_HUMAN	P29322 homo sapien
51	43.5	26.5	1799	1	LMB2_MOUSE	Q61292 mus musculus
52	43	26.2	63	1	MT_CHICK	P09376 gallus gall
53	43	26.2	75	1	DISC_TRIFL	P23323 trimeresuru
54	43	26.2	339	1	MBR1_YEAST	P23493 saccharomyc
55	43	26.2	494	1	ACH6_CHICK	P49581 gallus gall
56	43	26.2	540	1	ECM1_HUMAN	Q16610 homo sapien
57	43	26.2	571	1	RPSD_CHLMU	P56835 chlamydia m
58	43	26.2	571	1	RPSD_CHLTR	P18333 chlamydia t
59	43	26.2	572	1	RPSD_CHLPN	Q927f0 chlamydia p
60	43	26.2	754	1	PURL_MYCLE	Q50023 mycobacteri
61	43	26.2	837	1	UBPA_DICDI	P54201 dictyostelli
62	42.5	25.9	167	1	NIQK_KLEPN	P10392 klebsiella
63	42.5	25.9	305	1	YQBB_BACSU	P45918 bacillus su
64	42.5	25.9	581	1	IRR_RAT	Q64716 rattus norv
65	42.5	25.9	1032	1	AL4A_MOUSE	Q9w427 mus musculus
66	42.5	25.9	1353	1	PUR2_DROME	P00967 d trifuncti
67	42.5	25.9	1928	1	LPH_RAT	Q02401 rattus norv
68	42	25.6	106	1	COLA_HORSE	P02704 equus cabal
69	42	25.6	107	1	COL_RABIT	P42890 oryctolagus
70	42	25.6	183	1	Y389_HAEIN	P43991 haemophilus
71	42	25.6	227	1	AGI_ORYSA	P11219 oryza sativ
72	42	25.6	286	1	LEP4_BACNO	Q46525 bacteroides
73	42	25.6	353	1	ALC1_HUMAN	P01976 homo sapien
74	42	25.6	445	1	HH3R_HUMAN	O9y5n1 homo sapien
75	42	25.6	449	1	GATB_METTH	O27341 methanobact
76	42	25.6	506	1	ARSA_MOUSE	P50428 mus musculus
77	42	25.6	594	1	ZF37_MOUSE	P17141 mus musculus
78	42	25.6	602	1	YZ13_BORBU	Q60275 methanococc
79	42	25.6	631	1	RPSD_METJA	P52323 borrelia bu
80	42	25.6	679	1	PRTP_HSVSA	P24911 herpesvirus
81	42	25.6	855	1	ST14_HUMAN	Q9y5y6 homo sapien
82	42	25.6	903	1	AD12_MOUSE	Q61824 mus musculus
83	42	25.6	920	1	AD19_MOUSE	O35674 mus musculus
84	42	25.6	956	1	AD19_HUMAN	Q9h013 homo sapien
85	42	25.6	1151	1	KPC1_YEAST	P24583 saccharomyc
86	42	25.6	1944	1	CHD3_HUMAN	Q12873 homo sapien
87	41.5	25.3	416	1	ILIS_RAT	P43303 rattus norv
88	41.5	25.3	591	1	CO8B_HUMAN	P07358 homo sapien
89	41.5	25.3	814	1	AD15_HUMAN	Q14444 homo sapien
90	41.5	25.3	1712	1	NX2A_HUMAN	Q9p2s2 homo sapien
91	41.5	25.3	1715	1	NX2A_RAT	O63374 rattus norv
92	41.5	25.3	1835	1	CCAI_RAT	Q920y8 rattus norv
93	41.5	25.3	2871	1	FBN1_HUMAN	P35555 homo sapien
94	41	25.0	66	1	HST_YERKR	P31518 yersinia kr
95	41	25.0	94	1	DEFN_HUMAN	P11479 homo sapien
96	41	25.0	108	1	COLB_HORSE	P02705 equus cabal
97	41	25.0	422	1	HEM1_PSEAE	P42807 pseudomonas
98	41	25.0	445	1	HH3R_CAVPO	Q9j135 cavia porce
99	41	25.0	471	1	TRIA_BOVIN	O19131 bos taurus
100	41	25.0	559	1	PHO2_YEAST	P07269 saccharomyc

GenCore version 4.5.
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 07:47:17 ; Search time 51.98 Seconds
(without alignments)
61.969 Million cell updates/sec

Title: US-09-856-199-1

Perfect score: 164

Sequence: 1 KRCINQLLCKLPTDSELA PRSXCHCRHL 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

A_Geneseq_032802.*
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	161	98.2	29	21	AA031183 Human selenoprotein
2	161	98.2	103	21	AA031185 Human selenoprotein
3	161	98.2	381	21	AA031188 Human selenoprotein
4	111	67.7	21	21	AA031187 Human selenoprotein
5	111	67.7	110	21	AA037555 Human selenoprotein
6	111	67.7	240	21	AA057080 Human secreted protein
7	107	65.2	20	21	AA031186 Human selenoprotein
8	53.5	32.6	528	22	AA078837 Human protein SEQ
9	53.5	32.6	723	22	AA079821 Human protein SEQ
10	53	32.3	324	20	AA035364 Amino acid sequence
11	53	32.3	1460	21	AA040077 Arabidopsis thaliana

12	53	32.3	1462	21	AA040076 Arabidopsis thaliana
13	53	32.3	1518	21	AA040075 Arabidopsis thaliana
14	53	32.3	1604	21	AA031284 Arabidopsis thaliana
15	53	32.3	1606	21	AA031283 Arabidopsis thaliana
16	53	32.3	1662	21	AA031282 Arabidopsis thaliana
17	52	31.7	169	20	AA060558 Human normal bladder
18	52	31.7	363	22	AB059042 Drosophila melanogaster
19	52	31.7	430	20	AA031745 Mycobacterium tuberculosis
20	51	31.1	543	22	AA030622 Novel human secret
21	51	31.1	685	22	AB064268 Drosophila melanogaster
22	50	30.5	73	22	AA018623 Human lung antigen
23	50	30.5	121	21	AA058903 Arabidopsis thaliana
24	49.5	30.2	73	21	AA02921 Human secreted protein
25	49.5	30.2	2396	22	AB026979 Novel human diagnostic
26	49	29.9	1059	21	AA036839 Protein sequence analysis
27	49	29.9	1300	21	AA036838 Rat insulin receptor
28	48.5	29.6	59	21	AA028165 Arabidopsis thaliana
29	48.5	29.6	67	21	AA028164 Arabidopsis thaliana
30	48.5	29.6	132	20	AA023231 Erythrovirus v911
31	48.5	29.6	3080	10	AA093285 Sequence of clone
32	48.5	29.6	3210	9	AA081771 Deduced sequence of
33	48	29.3	129	21	AA071147 Pinctada fucata peptide
34	48	29.3	172	22	AA053277 Human gene 13 cDNA
35	48	29.3	296	21	AA054677 Arabidopsis thaliana
36	48	29.3	1268	21	AA036836 Human insulin receptor
37	48	29.3	1268	22	AB011702 Human insulin receptor
38	48	29.3	1297	21	AA036840 Human insulin receptor
39	47.5	29.0	70	22	AA058496 Propionibacterium
40	47.5	29.0	182	20	AA026136 Phospholipase A2
41	47.5	29.0	182	21	AA083611 Mature alpha chain
42	47.5	29.0	209	22	AB070477 Drosophila melanogaster
43	47.5	29.0	210	22	AA053377 Novel human diagnostic
44	47.5	29.0	1779	22	AB071649 Drosophila melanogaster
45	47	28.7	83	12	AA010111 Bitan-alpha. Trim
46	47	28.7	84	22	AB029847 Peptide #2498 cDNA
47	47	28.7	84	22	AB035027 Peptide #2533 cDNA
48	47	28.7	84	22	AB020438 Human brain expressed
49	47	28.7	84	22	AA055842 Human bone marrow
50	47	28.7	84	22	AA068215 Peptide #2571 cDNA
51	47	28.7	84	22	AA028534 Peptide #2446 cDNA
52	47	28.7	84	22	AA003764 Endosperm specific
53	47	28.7	124	19	AA056731 Human G protein-coupled
54	47	28.7	318	22	AA019243 Southern copperhead
55	47	28.7	483	21	AA079413 Agkistrodon contortrix
56	47	28.7	483	22	AA030336 Human protein sequence
57	47	28.7	505	22	AA092895 Human lung antigen
58	46.5	28.4	79	22	AA018598 Human colon cancer
59	46.5	28.4	426	22	AA073967 Human colon cancer
60	46	28.0	85	22	AA074885 Fragment of human
61	46	28.0	118	19	AA075244 Novel human secret
62	46	28.0	142	22	AA032800 Murine OR-like pol
63	46	28.0	317	22	AA072623 ADAM-8dis-Fc fusion
64	46	28.0	494	22	AA013051 Drosophila melanogaster
65	46	28.0	495	22	AB057748 Drosophila melanogaster
66	46	28.0	495	22	AB057752 Drosophila melanogaster
67	46	28.0	507	22	AB051183 Human sulfatase protein
68	46	28.0	660	22	AB068659 Drosophila melanogaster
69	46	28.0	817	22	AB066291 Drosophila melanogaster
70	46	28.0	824	22	AA020251 Human ADAM8. Homo
71	45.5	27.7	177	20	AA070007 Breast cancer associated
72	45.5	27.7	182	20	AA026110 Phospholipase A2
73	45.5	27.7	182	20	AA026111 Phospholipase A2
74	45.5	27.7	183	20	AA026112 Phospholipase A2
75	45.5	27.7	202	21	AA083648 NSI Phospholipase
76	45.5	27.7	218	22	AA074312 Human colon cancer
77	45.5	27.7	1017	22	AB064674 Drosophila melanogaster
78	45.5	27.7	1801	19	AA050895 Rat laminin B2 chain
79	45	27.4	82	15	AA053933 Disintegrin peptide
80	45	27.4	83	10	AA091090 Sequence of viper
81	45	27.4	83	11	AA06391 Bitistatin 1. Bitis
82	45	27.4	83	12	AA012988 Bitistatin 1 fibrin
83	45	27.4	84	21	AA041670 Human ORFX ORF1434
84	45	27.4	96	22	AB012841 Novel human diagnostic

85 Bacteriophage Dp-1
86 Human digestive sy
87 Human liver associ
88 Arabidopsis thalia
89 Arabidopsis thalia
90 Novel human secret
91 Human p53 target m
92 Amino acid sequenc
93 Arabidopsis thalia
94 Arabidopsis thalia
95 Arabidopsis thalia
96 Arabidopsis thalia
97 Arabidopsis thalia
98 Arabidopsis thalia
99 Amino acid sequenc
100 Drosophila melanog

ALIGNMENTS

RESULT 1
AAB03183
ID AAB03183 standard; peptide; 29 AA.
XX
AC AAB03183;
XX
DT 23-OCT-2000 (first entry)
XX
DE Human selenoprotein P C-terminal fragment, SEQ ID NO:1.
XX

XX Selenoprotein P fragment; human; programmed cell death inhibitor;
KW apoptosis inhibitor; antiapoptotic; acquired immunodeficiency syndrome;
KW AIDS; Parkinson's disease; Alzheimer's disease; myocardial infarct;
KW cerebral infarct; stroke; reperfusion injury; organ transplant rejection;
KW arteriosclerosis; redox disorder; immune disorder; cardiovascular;
KW cerebroprotective; culture medium additive.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 22 /note= "Selenocysteine"
FT
XX WO200031131-A1.
XX
PN 02-JUN-2000.
XX
PD 12-NOV-1999; 99WO-JP06322.
XX
PF 19-NOV-1998; 98JP-0347863.
XX
PR (KAGA) CHEMO-SERO-THERAPEUTIC RES INST.
XX
PA Hirashima M, Maeda H, Nozaki C;
XX
XX WPI; 2000-451677/39.
XX
DR Peptide fragments with cell death inhibitory activity, useful in
PT preventing and treating apoptosis-associated diseases particularly
PT caused by stress e.g. Parkinson's disease, Alzheimer's and
PT arteriosclerosis
XX
XX Claim 2; Page 31; 56pp; Japanese.
XX
XX This sequence represents a fragment of human selenoprotein P which
CC has anti-apoptotic activity. The invention relates to fragments
CC (AAB03183, AAB03184) derived from the C-terminal 103 amino acids of
CC human selenoprotein P (AAB03186) which act as inhibitors of apoptosis
CC (programmed cell death). Their use in therapeutic compositions and cell
CC culture media, and antibodies against them. The invention also relates to
CC a method for studying the phenomenon of sudden cell death in a human
CC megakaryotic series gemmule cell culture system involving use of the

CC anti-apoptotic peptides of the invention. Selenoprotein P is a plasma
CC protein which confers protection against peroxynitrite-mediated
CC oxidation and nitration. The peptide fragments can be used as
CC prophylactic and therapeutic agents for apoptosis-associated diseases,
CC including AIDS (acquired immunodeficiency syndrome), Parkinson's disease,
CC Alzheimer's disease, arteriosclerosis, myocardial infarct, cerebral
CC infarct, organ transplant rejection, and reperfusion injury. They may
CC also be used in the prevention or treatment of disorders relating to
CC redox reactions, or of disorders of the immune system. They may
CC additionally be used as an additive for cell culture and for screening
CC cell death activity in vitro.
XX
SQ Sequence 29 AA:

Query Match 98.2%; Score 161; DB 21; Length 29;
Best Local Similarity 96.6%; Pred. No. 1.7e-14;
Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KRCINOLLCKLPDSELA PRSXCHCHRL 29
|||||
Db 1 krcingllcklpdtselaprsccchchrl 29

RESULT 2
AAB03185
ID AAB03185 standard; peptide; 103 AA.
XX
AC AAB03185;
XX
DT 23-OCT-2000 (first entry)
XX
DE Human selenoprotein P C-terminal 103 residues.
XX

XX Selenoprotein P fragment; human; programmed cell death inhibitor;
KW apoptosis inhibitor; antiapoptotic; acquired immunodeficiency syndrome;
KW AIDS; Parkinson's disease; Alzheimer's disease; myocardial infarct;
KW cerebral infarct; stroke; reperfusion injury; organ transplant rejection;
KW arteriosclerosis; redox disorder; immune disorder; cardiovascular;
KW cerebroprotective; culture medium additive.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 22 /note= "Selenocysteine"
FT
FT Modified-site 40 /note= "Selenocysteine"
FT
FT Modified-site 52 /note= "Selenocysteine"
FT
FT Modified-site 67 /note= "Selenocysteine"
FT
FT Modified-site 74 /note= "Selenocysteine"
FT
FT Modified-site 89 /note= "Selenocysteine"
FT
FT Modified-site 91 /note= "Selenocysteine"
FT
FT Modified-site 98 /note= "Selenocysteine"
FT
FT Modified-site 100 /note= "Selenocysteine"
FT
FT
XX WO200031131-A1.
XX
PN 02-JUN-2000.
XX
PD 12-NOV-1999; 99WO-JP06322.
XX
PF 19-NOV-1998; 98JP-0347863.
XX
PR (KAGA) CHEMO-SERO-THERAPEUTIC RES INST.
XX
PA

PI Hirashima M, Maeda H, Nozaki C;
 DR WPI; 2000-451677/39.
 XX
 XX Peptide fragments with cell death inhibitory activity, useful in
 PT preventing and treating apoptosis-associated diseases particularly
 PT caused by stress e.g. Parkinson's disease, Alzheimer's and
 PT arteriosclerosis
 XX
 PS Disclosure; Page 11; 56pp; Japanese.
 XX
 CC The invention relates to fragments (AAB03183, AAB03184) derived from the
 CC C-terminal 103 amino acids of human selenoprotein P (AAB03188) which act
 CC as inhibitors of apoptosis (programmed cell death), their use in
 CC therapeutic compositions and cell culture media, and antibodies against
 CC them. The invention also relates to a method for studying the phenomenon
 CC of sudden cell death in a human megakaryotic series gemmule cell culture
 CC system involving use of the anti-apoptotic peptides of the invention.
 CC Selenoprotein P is a plasma protein which confers protection against
 CC peroxynitrite-mediated oxidation and nitration. The peptide fragments
 CC can be used as prophylactic and therapeutic agents for
 CC apoptosis-associated diseases, including AIDS (acquired immunodeficiency
 CC syndrome), Parkinson's disease, Alzheimer's disease, arteriosclerosis,
 CC myocardial infarct, cerebral infarct, organ transplant rejection, and
 CC reperfusion injury. They may also be used in the prevention or treatment
 CC of disorders relating to redox reactions, or of disorders of the immune
 CC system. They may additionally be used as an additive for cell culture and
 CC for screening cell death activity in vitro. The present sequence
 CC represents the C-terminal 103 amino acids of human selenoprotein P.
 XX
 SQ Sequence 103 AA;
 Query Match 98.2%; Score 161; DB 21; Length 103;
 Best Local Similarity 96.6%; Pred. No. 5.7e-14;
 Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 KRCINQLLCKLPTDSELAPRSXCCHRH 29
 DB 1 krcinqllicklptdse laprscchchrhl 29
 RESULT 3
 AAB03188
 ID AAB03188 standard; peptide; 381 AA.
 XX
 AC AAB03188;
 XX
 DT 23-OCT-2000 (first entry)
 DE Human selenoprotein P.
 KW Selenoprotein P; human; programmed cell death inhibitor;
 KW apoptosis inhibitor; antiapoptotic; acquired immunodeficiency syndrome;
 KW AIDS; Parkinson's disease; Alzheimer's disease; myocardial infarct;
 KW cerebral infarct; stroke; reperfusion injury; organ transplant rejection;
 KW arteriosclerosis; redox disorder; immune disorder; cardiovascular;
 KW cerebroprotective; culture medium additive.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..19
 FT /note= "Signal peptide"
 FT Protein 19..381
 FT /note= "Mature human selenoprotein P"
 FT Modified-site 281
 FT /note= "Selenocysteine"
 FT Modified-site 299
 FT /note= "Selenocysteine"
 FT Modified-site 311
 FT /note= "Selenocysteine"
 FT Modified-site 326

FT Modified-site /note= "Selenocysteine"
 FT 333
 FT /note= "Selenocysteine"
 FT 348
 FT /note= "Selenocysteine"
 FT 350
 FT /note= "Selenocysteine"
 FT 357
 FT /note= "Selenocysteine"
 FT 359
 FT /note= "Selenocysteine"
 XX
 XX WO200031131-A1.
 PN 02-JUN-2000.
 PD 12-NOV-1999; 99WO-JP06322.
 XX 19-NOV-1998; 98JP-0347863.
 PR (KAGA) CHEMO-SERO-THERAPEUTIC RES INST.
 XX Hirashima M, Maeda H, Nozaki C;
 WPI; 2000-451677/39.
 XX
 PT Peptide fragments with cell death inhibitory activity, useful in
 PT preventing and treating apoptosis-associated diseases particularly
 PT caused by stress e.g. Parkinson's disease, Alzheimer's and
 PT arteriosclerosis
 XX
 PS Example 3; Page 20; 56pp; Japanese.
 XX
 CC The invention relates to fragments (AAB03183, AAB03184) derived from the
 CC C-terminal 103 amino acids of human selenoprotein P (AAB03188) which act
 CC as inhibitors of apoptosis (programmed cell death), their use in
 CC therapeutic compositions and cell culture media, and antibodies against
 CC them. The invention also relates to a method for studying the phenomenon
 CC of sudden cell death in a human megakaryotic series gemmule cell culture
 CC system involving use of the anti-apoptotic peptides of the invention.
 CC Selenoprotein P is a plasma protein which confers protection against
 CC peroxynitrite-mediated oxidation and nitration. The peptide fragments
 CC can be used as prophylactic and therapeutic agents for
 CC apoptosis-associated diseases, including AIDS (acquired immunodeficiency
 CC syndrome), Parkinson's disease, Alzheimer's disease, arteriosclerosis,
 CC myocardial infarct, cerebral infarct, organ transplant rejection, and
 CC reperfusion injury. They may also be used in the prevention or treatment
 CC of disorders relating to redox reactions, or of disorders of the immune
 CC system. They may additionally be used as an additive for cell culture and
 CC for screening cell death activity in vitro. The present sequence
 CC represents human selenoprotein P.
 XX
 SQ Sequence 381 AA;
 Query Match 98.2%; Score 161; DB 21; Length 381;
 Best Local Similarity 96.6%; Pred. No. 1.9e-13;
 Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 KRCINQLLCKLPTDSELAPRSXCCHRH 29
 DB 279 krcinqllicklptdse laprscchchrhl 307
 RESULT 4
 AAB03187
 ID AAB03187 standard; peptide; 21 AA.
 XX
 AC AAB03187;
 XX
 DT 23-OCT-2000 (first entry)
 DE Human selenoprotein P C-terminal fragment, SEQ ID NO:5.

XX Selenoprotein P fragment; human; programmed cell death inhibitor;
 KW apoptosis inhibitor; antiapoptotic; acquired immunodeficiency syndrome;
 KW AIDS; Parkinson's disease; Alzheimer's disease; myocardial infarct;
 KW cerebral infarct; stroke; reperfusion injury; organ transplant rejection;
 KW arteriosclerosis; redox disorder; immune disorder; cardiovascular;
 KW cerebroprotective; culture medium additive.

XX Homo sapiens.
 XX WO200031131-A1.
 PN
 XX
 XX 02-JUN-2000.
 PD
 XX
 XX 12-NOV-1999; 99WO-JP06322.
 PF
 XX
 XX 19-NOV-1998; 98JP-0347863.
 PR
 XX
 XX (KAGA) CHEMO-SERO-THERAPEUTIC RES INST.
 PA

XX Hirashima M, Maeda H, Nozaki C;
 PI
 XX
 XX WPI; 2000-451677/39.
 DR
 XX

PT Peptide fragments with cell death inhibitory activity, useful in
 PT preventing and treating apoptosis-associated diseases particularly
 PT caused by stress e.g. Parkinson's disease, Alzheimer's and
 PT arteriosclerosis -

PS Example 3; Page 19; 56pp; Japanese.

XX The invention relates to fragments (AAB03183, AAB03184) derived from the
 CC C-terminal 103 amino acids of human selenoprotein P (AAB03188) which act
 CC as inhibitors of apoptosis (programmed cell death), their use in
 CC therapeutic compositions and cell culture media, and antibodies against
 CC them. The invention also relates to a method for studying the phenomenon
 CC of sudden cell death in a human megakaryotic series gemmule cell culture
 CC system involving use of the anti-apoptotic peptides of the invention.
 CC Selenoprotein P is a plasma protein which confers protection against
 CC peroxynitrite-mediated oxidation and nitration. The peptide fragments
 CC can be used as prophylactic and therapeutic agents for
 CC apoptosis-associated diseases, including AIDS (acquired immunodeficiency
 CC syndrome), Parkinson's disease, Alzheimer's disease, arteriosclerosis,
 CC myocardial infarct, cerebral infarct, organ transplant rejection, and
 CC reperfusion injury. They may also be used in the prevention or treatment
 CC of disorders relating to redox reactions, or of disorders of the immune
 CC system. They may additionally be used as an additive for cell culture and
 CC for screening cell death activity in vitro. The present sequence
 CC represents a C-terminal fragment of human selenoprotein P used in an
 CC exemplification of the invention.

XX Sequence 21 AA;

Query Match 67.7%; Score 111; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 5.1e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRCINQLLCKLPTDSELAPRS 21
 Db 1 krcinqllickiptdselaprs 21

RESULT 5
 AAG03755
 ID AAG03755 standard; Protein; 110 AA.
 AC AAG03755;
 XX

XX 06-OCT-2000 (first entry)
 XX Human secreted protein, SEQ ID NO: 7836.
 DE
 XX

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.

XX Homo sapiens.

XX EPI033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX N-PSDB; AAC03761.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 13; SEQ ID 7836; 71pp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.

XX Sequence 110 AA;

Query Match 67.7%; Score 111; DB 21; Length 110;
 Best Local Similarity 100.0%; Pred. No. 2.4e-07;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRCINQLLCKLPTDSELAPRS 21

Db 90 krcinqllickiptdselaprs 110

RESULT 6
 AAB57080
 ID AAB57080 standard; Protein; 240 AA.
 XX
 XX AAB57080;
 XX
 XX 13-MAR-2001 (first entry)
 XX
 XX Human prostate cancer antigen protein sequence SEQ ID NO:1658.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
 KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
 KW vulnery; gastrointestinal; nephrotropic; antineoplastic; gynaecological;
 KW antibacterial; gene therapy; immune; immune; reproductive; renal;
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
 KW wound; infectious disease.

XX Homo sapiens.

XX WO200055174-A1.

XX


```

PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05988.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587513/55.
DR N-PSDB; AAF16283.
XX
PT Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer.
XX
PS Claim 11; Page 2116-2117; 2338pp; English.
XX
CC AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB5363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioactive, immunomodulatory, muscular, vulnery, gastrointestinal,
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 240 AA;

Query Match 67.7%; Score 111; DB 21; Length 240;
Best Local Similarity 100.0%; Pred. No. 5.1e-07;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KRCINOLCKLPTDSELAPRS 21
| | | | | | | | | | | | | | | | | | | | |
Db 220 krcinqlcklptdselapr 240

RESULT 7
AAB03186
ID AAB03186 standard; peptide; 20 AA.
XX
AC AAB03186;
XX
DT 23-OCT-2000 (first entry)
XX
DE Human selenoprotein P C-terminal fragment, SEQ ID NO:4.
XX
KW Selenoprotein P fragment; human; programmed cell death inhibitor;
KW apoptosis inhibitor; antiapoptotic; acquired immunodeficiency syndrome;
KW AIDS; Parkinson's disease; Alzheimer's disease; myocardial infarct;
KW cerebral infarct; stroke; reperfusion injury; organ transplant rejection;
KW arteriosclerosis; redox disorder; immune disorder; cardiovascular;
KW cerebroprotective; culture medium additive.
XX
OS Homo sapiens.
XX
PN WO200031131-A1.
XX
PD 02-JUN-2000.
XX
PF 12-NOV-1999; 99WO-JP06322.
XX
PR 19-NOV-1998; 98JP-0347863.

(KAGA ) CHEMO-SERO-THERAPEUTIC RES INST.
Hirashima M, Maeda H, Nozaki C;
WPI; 2000-451677/39.
Peptide fragments with cell death inhibitory activity, useful in
preventing and treating apoptosis-associated diseases particularly
arteriosclerosis.
Example 4; Page 11; 56pp; Japanese.
The invention relates to fragments (AAB03183, AAB03184) derived from the
C-terminal 103 amino acids of human selenoprotein P (AAB03188) which act
as inhibitors of apoptosis (programmed cell death), their use in
therapeutic compositions and cell culture media, and antibodies against
them. The invention also relates to a method for studying the phenomenon
of sudden cell death in a human megakaryotic series gemmule cell culture
system involving use of the anti-apoptotic peptides of the invention.
Selenoprotein P is a plasma protein which confers protection against
peroxynitrite-mediated oxidation and nitration. The peptide fragments
can be used as prophylactic and therapeutic agents for
apoptosis-associated diseases, including AIDS (acquired immunodeficiency
syndrome), Parkinson's disease, Alzheimer's disease, arteriosclerosis,
myocardial infarct, cerebral infarct, organ transplant rejection, and
reperfusion injury. They may also be used in the prevention or treatment
of disorders relating to redox reactions, or of disorders of the immune
system. They may additionally be used as an additive for cell culture and
for screening cell death activity in vitro. The present sequence
represents a C-terminal fragment of human selenoprotein P used in an
exemplification of the invention.
XX
SQ Sequence 20 AA;

Query Match 65.2%; Score 107; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KRCINOLCKLPTDSELAPR 20
| | | | | | | | | | | | | | | | | | | | |
Db 1 krcinqlcklptdselapr 20

RESULT 8
AAM78837
ID AAM78837 standard; Protein; 528 AA.
XX
AC AAM78837;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 1499.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.

```

```

PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX WPI; 2001-476283/51.
DR N-PSDB; AAK51970.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
PT
XX
PS Claim 20; Page 3789-3790; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 528 AA;

Query Match 32.6%; Score 53.5; DB 22; Length 528;
Best Local Similarity 38.5%; Pred. NO. 41;
Matches 10; Conservative 5; Mismatches 8; Indels 3; Gaps 1;

QY 1 KRCINQLLCKLPTDSELA PRSXCHC 26
   |:| |:| |:| ::| |:| |:|
Db 10 Kqcknlltckv---aqvcpcgclhc 32

RESULT 9
AAM79821
ID AAM79821 standard; Protein; 723 AA.
XX
AC AAM79821;
XX
XX 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 3467.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
XX WO200157190-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US04098.
XX
XX 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR '19-JUL-2000; 2000US-0620325.

```

```

PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX WPI; 2001-476283/51.
DR N-PSDB; AAK52954.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
PT
XX
PS Claim 20; Page 347; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 723 AA;

Query Match 32.6%; Score 53.5; DB 22; Length 723;
Best Local Similarity 38.5%; Pred. NO. 56;
Matches 10; Conservative 5; Mismatches 8; Indels 3; Gaps 1;

QY 1 KRCINQLLCKLPTDSELA PRSXCHC 26
   |:| |:| |:| ::| |:| |:|
Db 10 Kqcknlltckv---aqvcpcgclhc 32

RESULT 10
AAY35364
ID AAY35364 standard; Protein; 324 AA.
XX
AC AAY35364;
XX
XX 13-SEP-1999 (first entry)
XX
DE Amino acid sequence of a Chlamydia pneumoniae protein.
XX
XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW vaccine; neutralising epitope.
XX
OS Chlamydia pneumoniae.
XX
XX WO9927105-A2.
XX
XX 03-JUN-1999.
XX
XX 20-NOV-1998; 98WO-IB01890.
PR 04-NOV-1998; 98US-0107078.
PR 21-NOV-1997; 97ER-0014673.
XX
XX (GEST ) GENSET.
XX

```

PI Griffals R;
XX WPI; 1999-357842/30.
XX
XX Genome sequence of Chlamydia pneumoniae
XX
PS Page 1160; Disclosure; 1912pp; English.
XX
CC AAY34584-Y35879 represent the proteins encoded by all the open reading
CC frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.
CC C. pneumoniae causes respiratory disease such as pneumonia and
CC bronchitis and is thought to be a contributing factor in heart
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC nodosum or pharyngitis. The polypeptides encoded by the open reading
CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC nucleotide sequences can also be used as immunogenic compositions,
CC especially where the vector directs the expression of a neutralising
CC epitope of C. pneumoniae.
XX
SQ Sequence 324 AA;

Query Match 32.3%; Score 53; DB 20; Length 324;
Best Local Similarity 46.7%; Pred. No. 30;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 13 TDSELPASXCCHCR 27
:|::||:||||:
Db 236 sdpdvapiaecchcq 250

RESULT 11
AAG40077
ID AAG40077 standard; Protein; 1460 AA.
XX
AC AAG40077;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 49678.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 99US-0132485.

PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.

```
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 29-SEP-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 32.3%; Score 53; DB 21; Length 1460;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 KLPTDSELAPRSXCCH 25
   ||| ||:| |||
Db 485 rlpedsdvamedacch 500

RESULT 12
AAG40076
ID AAG40076 standard; Protein; 1462 AA.
XX
AC AAG40076;
DT 18-OCT-2000 (first entry)
DE
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 49677.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
```

PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 22-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

```
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 32.3%; Score 53; DB 21; Length 1462;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 KLPTDSELAPRSXCCH 25
Db 487 rlpedsdvamedacch 502

RESULT 13
AAG40075
ID AAG40075 standard; Protein: 1518 AA.
XX
AC AAG40075;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 49676.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132488.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134376.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
```

PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 01-SEP-1999; 99US-0152363.
 PR 07-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 13-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0156559.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160584.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.

PR 29-OCT-1999; 99US-0162142.
 Query Match 32.3%; Score 53; DB 21; Length 1518;
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 10 KLPTDSELA PRSCXCH 25
 :||| ||:| |
 Db 543 rlpedsdvamedacch 558
 RESULT 14
 AAG31284
 ID AAG31284 standard; Protein; 1604 AA.
 XX AC AAG31284;
 XX DT 17-OCT-2000 (first entry)
 XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 37543.
 XX KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX OS Arabidopsis thaliana.
 XX PN EP1033405-A2.
 XX PD 06-SEP-2000.
 XX PF 25-FEB-2000; 2000EP-0301439.
 XX PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 23-MAR-1999; 99US-0123548.
 PR 25-MAR-1999; 99US-0125788.
 PR 29-MAR-1999; 99US-0126264.
 PR 01-APR-1999; 99US-0126785.
 PR 06-APR-1999; 99US-0127462.
 PR 08-APR-1999; 99US-0128234.
 PR 16-APR-1999; 99US-0128714.
 PR 19-APR-1999; 99US-0129845.
 PR 21-APR-1999; 99US-0130077.
 PR 23-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 20-MAY-1999; 99US-0134941.
 PR 21-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 03-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 32.3%; Score 53; DB 21; Length 1604;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 KLPTDSELAPRSXCCH 25


```
Db      485  ripedsdvamedacch 500
      :|| ||::| |||
RESULT 15
AAG31283
ID  AAG31283 standard; Protein; 1606 AA.
XX
AC  AAG31283;
XX
DT  17-OCT-2000 (first entry)
XX
DE  Arabidopsis thaliana protein fragment SEQ ID NO: 37542.
XX
KW  Protein identification; signal transduction pathway; metabolic pathway;
KW  hybridisation assay; genetic mapping; gene expression control; promoter;
KW  termination sequence.
XX
OS  Arabidopsis thaliana.
XX
PN  EP1033405-A2.
XX
PD  06-SEP-2000.
XX
PF  25-FEB-2000; 2000EP-0301439.
XX
PR  25-FEB-1999; 99US-0121825.
PR  05-MAR-1999; 99US-0123180.
PR  09-MAR-1999; 99US-0123548.
PR  23-MAR-1999; 99US-0125788.
PR  25-MAR-1999; 99US-0126264.
PR  29-MAR-1999; 99US-0126785.
PR  01-APR-1999; 99US-0127462.
PR  06-APR-1999; 99US-0128234.
PR  08-APR-1999; 99US-0128714.
PR  16-APR-1999; 99US-0129845.
PR  19-APR-1999; 99US-0130077.
PR  21-APR-1999; 99US-0130449.
PR  23-APR-1999; 99US-0130510.
PR  28-APR-1999; 99US-0130891.
PR  30-APR-1999; 99US-0131449.
PR  30-APR-1999; 99US-0132048.
PR  04-MAY-1999; 99US-0132407.
PR  05-MAY-1999; 99US-0132484.
PR  06-MAY-1999; 99US-0132485.
PR  07-MAY-1999; 99US-0132486.
PR  11-MAY-1999; 99US-0132487.
PR  14-MAY-1999; 99US-0132863.
PR  14-MAY-1999; 99US-0134218.
PR  14-MAY-1999; 99US-0134219.
PR  14-MAY-1999; 99US-0134221.
PR  18-MAY-1999; 99US-0134370.
PR  19-MAY-1999; 99US-0134768.
PR  20-MAY-1999; 99US-0134941.
PR  21-MAY-1999; 99US-0135124.
PR  24-MAY-1999; 99US-0135353.
PR  25-MAY-1999; 99US-0135629.
PR  27-MAY-1999; 99US-0136021.
PR  28-MAY-1999; 99US-0136392.
PR  01-JUN-1999; 99US-0136782.
PR  03-JUN-1999; 99US-0137222.
PR  04-JUN-1999; 99US-0137528.
PR  07-JUN-1999; 99US-0137502.
PR  08-JUN-1999; 99US-0137724.
PR  10-JUN-1999; 99US-0138094.
PR  10-JUN-1999; 99US-0138540.
PR  14-JUN-1999; 99US-0138847.
PR  16-JUN-1999; 99US-0139119.
PR  16-JUN-1999; 99US-0139452.
PR  17-JUN-1999; 99US-0139453.
PR  17-JUN-1999; 99US-0139492.
PR  18-JUN-1999; 99US-0139454.
PR  18-JUN-1999; 99US-0139455.
PR  18-JUN-1999; 99US-0139456.
PR  18-JUN-1999; 99US-0139457.
PR  18-JUN-1999; 99US-0139458.
PR  18-JUN-1999; 99US-0139459.
PR  18-JUN-1999; 99US-0139460.
PR  18-JUN-1999; 99US-0139461.
PR  18-JUN-1999; 99US-0139462.
PR  18-JUN-1999; 99US-0139463.
PR  18-JUN-1999; 99US-0139750.
PR  18-JUN-1999; 99US-0139763.
PR  21-JUN-1999; 99US-0139817.
PR  22-JUN-1999; 99US-0139899.
PR  23-JUN-1999; 99US-0140353.
PR  23-JUN-1999; 99US-0140354.
PR  24-JUN-1999; 99US-0140695.
PR  28-JUN-1999; 99US-0140823.
PR  29-JUN-1999; 99US-0140991.
PR  30-JUN-1999; 99US-0141287.
PR  01-JUL-1999; 99US-0141842.
PR  01-JUL-1999; 99US-0142154.
PR  02-JUL-1999; 99US-0142055.
PR  06-JUL-1999; 99US-0142390.
PR  08-JUL-1999; 99US-0142803.
PR  09-JUL-1999; 99US-0142920.
PR  12-JUL-1999; 99US-0142977.
PR  13-JUL-1999; 99US-0143542.
PR  14-JUL-1999; 99US-0143624.
PR  15-JUL-1999; 99US-0144005.
PR  16-JUL-1999; 99US-0144085.
PR  16-JUL-1999; 99US-0144086.
PR  19-JUL-1999; 99US-0144325.
PR  19-JUL-1999; 99US-0144331.
PR  19-JUL-1999; 99US-0144332.
PR  19-JUL-1999; 99US-0144333.
PR  19-JUL-1999; 99US-0144334.
PR  19-JUL-1999; 99US-0144335.
PR  20-JUL-1999; 99US-0144352.
PR  20-JUL-1999; 99US-0144632.
PR  20-JUL-1999; 99US-0144884.
PR  21-JUL-1999; 99US-0144814.
PR  21-JUL-1999; 99US-0145086.
PR  21-JUL-1999; 99US-0145088.
PR  22-JUL-1999; 99US-0145085.
PR  22-JUL-1999; 99US-0145087.
PR  22-JUL-1999; 99US-0145089.
PR  22-JUL-1999; 99US-0145192.
PR  23-JUL-1999; 99US-0145145.
PR  23-JUL-1999; 99US-0145218.
PR  23-JUL-1999; 99US-0145224.
PR  26-JUL-1999; 99US-0145276.
PR  27-JUL-1999; 99US-0145913.
PR  27-JUL-1999; 99US-0145918.
PR  27-JUL-1999; 99US-0145919.
PR  28-JUL-1999; 99US-0145951.
PR  02-AUG-1999; 99US-0146386.
PR  02-AUG-1999; 99US-0146388.
PR  02-AUG-1999; 99US-0146389.
PR  03-AUG-1999; 99US-0147038.
PR  04-AUG-1999; 99US-0147204.
PR  04-AUG-1999; 99US-0147302.
PR  05-AUG-1999; 99US-0147192.
PR  05-AUG-1999; 99US-0147260.
PR  06-AUG-1999; 99US-0147303.
PR  06-AUG-1999; 99US-0147416.
PR  09-AUG-1999; 99US-0147493.
PR  09-AUG-1999; 99US-0147935.
PR  10-AUG-1999; 99US-0148171.
PR  11-AUG-1999; 99US-0148319.
PR  12-AUG-1999; 99US-0148341.
PR  13-AUG-1999; 99US-0148565.
PR  13-AUG-1999; 99US-0148684.
PR  16-AUG-1999; 99US-0149368.
```

```
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      32.3%; Score 53; DB 21; Length 1606;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches      8; Conservative      3; Mismatches      5; Indels      0; Gaps      0;

QY 10 KLPDTDELAPRSXCCH 25
   :||| ||::| |||
Db 487 rlpedsdvamedacch 502

RESULT 16
RAG31282
ID RAG31282 standard; Protein: 1662 AA.
```

```
XX AAG31282;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 37541.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
```

PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.

PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 32.3%; Score 53; DB 21; Length 1662;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 KLPTDSELAPRSXCCH 25
:|:|:|:|:|

Db 543 rlpedsdvamedacch 558

RESULT 17

AY60558
ID AAY60558 standard; Protein; 169 AA.
XX

AC AAY60558;

XX 31-JAN-2000 (first entry)

XX Human normal bladder tissue EST encoded protein 230.
DE
XX


```
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232403.
PR 14-SEP-2000; 2000US-0232406.
PR 14-SEP-2000; 2000US-0232423.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 26-SEP-2000; 2000US-0234998.
PR 27-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246533.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.

PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-457723/49.
XX N-PSDB; AAS29910.
XX
PT Isolated polypeptide for treating, preventing and/or prognosing
PT respiratory disorders related to the lung including lung cancers and
PT also for testing and detection e.g. diagnosis
XX
XX Claim 11; SEQ ID No 174; 507pp; English.
XX
XX Sequences AAS29839-AAS29930 represent the lung antigen polypeptides of
XX the invention. Lung antigen polypeptides and their associated
XX polynucleotides are useful in the diagnosis, treatment and prevention of
XX various types of disorders in e.g. humans, mice, rabbits, goats, horses,
XX cats, dogs, chickens or sheep. A pathological condition can be determined
XX by detecting the presence or absence of a mutation in a lung antigen
XX polynucleotide. The treatable disorders include autoimmune diseases such
XX as rheumatoid arthritis, hyperproliferative disorders such as neoplasms
XX of the breast or liver, cardiovascular disorders such as cardiac arrest,
XX cerebrovascular disorders such as cerebral ischaemia, nervous system
XX disorders such as Alzheimer's disease, infections caused by bacteria,
XX viruses and fungi, ocular disorders such as corneal infection, endocrine
XX disorders such as premature labour, gastrointestinal disorders such as
XX Crohn's disease, renal disorders such as glomerulonephritis and
XX respiratory disorders such as asthma and pleurisy. The polypeptides can
XX also be used to aid wound healing, to prevent skin aging due to sunburn,
XX to maintain organs before transplantation, to regenerate tissues and in
XX chemotaxis. The polypeptides can also be used as a food additive or
XX preservative to increase or decrease storage capabilities.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

Query Match 30.5%; Score 50; DB 22; Length 73;
Best Local Similarity 47.4%; Pred. No. 19;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 8 LCKLPTDSELAPRSXCCHC 26
Db 45 lcelppgphlpggsllcsc 63
||:| | | | |
| | | | | |

RESULT 23
AAG58903
ID AAG58903 standard; Protein; 121 AA.
XX
XX AAG58903;
XX
XX 18-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 76126.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
```

OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD
XX PF 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PF 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132866.
PR 11-MAY-1999; 99US-0134258.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.

PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID No 57338; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 2396 AA;

Query Match 30.2%; Score 49.5; DB 22; Length 2396;
 Best Local Similarity 40.5%; Pred. No. 5.8e+02;
 Matches 15; Conservative 2; Mismatches 9; Indels 11; Gaps 3;

QY 2 RCINQLLCKLPTD-----SELAPRXCCH-CRHL 29
 ||||| ||| ||: || | ||:
 Db 1498 rcin--taklpkxvvalfshpgelsarslhaacrh 1532

RESULT 26
 AAB36839
 ID AAB36839 standard; Protein; 1069 AA.

XX
 AC AAB36839;
 XX
 DT 20-FEB-2001 (first entry)
 XX
 DE Protein sequence associated with IRRR.

XX
 KW Insulin receptor-related receptor; IRRR; chromosome 1q21-q24;
 KW obesity; dyslipidemia; diabetes.

OS Rattus rattus.

XX WO2000065090-A2.

XX 02-NOV-2000.

XX 19-APR-2000; 2000WO-US10644.

XX 22-APR-1999; 99US-0296906.

XX 22-JUN-1999; 99US-0337976.

XX (ZYMO) ZYMOGENETICS INC.

XX Lok S, Whitmore TE;

XX WPI; 2000-687365/67.

XX Detecting a chromosome 1q21-q24 abnormality for diagnosing metabolic
 PT disease, such as human obesity and diabetic disorders, comprises
 PT examining insulin receptor-related receptor gene and its gene products

PT -
 XX Claim 21; Page 94-97; 111pp; English.
 XX
 CC The present invention relates to insulin receptor-related receptor
 CC (IRRR). Mutations in this gene indicate a chromosome 1q21-q24
 CC abnormality. IRRR polypeptides and DNA may be useful in the diagnosis of
 CC of disorders associated with abnormal expression of the IRRR protein,
 CC for example obesity, dyslipidemia and diabetes.
 XX
 SQ Sequence 1069 AA;

Query Match 29.9%; Score 49; DB 21; Length 1069;
 Best Local Similarity 32.3%; Pred. No. 3.2e+02;
 Matches 10; Conservative 2; Mismatches 7; Indels 12; Gaps 1;

QY 10 KLPTDS-----ELAPRXCCHCRH 28
 :||| | | | || | |
 Db 721 rlptsshdtrfdredpaleaepeggcpcqh 751

RESULT 27

AAB36838
 ID AAB36838 standard; Protein; 1300 AA.

XX AAB36838;

XX 20-FEB-2001 (first entry)

XX Rat Insulin receptor-related receptor protein.

XX Insulin receptor-related receptor; IRRR; chromosome 1q21-q24;
 KW obesity; dyslipidemia; diabetes.

OS Rattus rattus.

XX WO2000065090-A2.

XX 02-NOV-2000.

XX 19-APR-2000; 2000WO-US10644.

XX 22-APR-1999; 99US-0296906.

XX 22-JUN-1999; 99US-0337976.

XX (ZYMO) ZYMOGENETICS INC.

XX Lok S, Whitmore TE;

XX WPI; 2000-687365/67.

XX N-PSDB; AAC68342.

XX Detecting a chromosome 1q21-q24 abnormality for diagnosing metabolic
 PT disease, such as human obesity and diabetic disorders, comprises
 PT examining insulin receptor-related receptor gene and its gene products

XX Claim 11; Page 82-85; 111pp; English.

XX The present invention relates to insulin receptor-related receptor
 CC (IRRR). Mutations in this gene indicate a chromosome 1q21-q24
 CC abnormality. IRRR polypeptides and DNA may be useful in the diagnosis of
 CC of disorders associated with abnormal expression of the IRRR protein,
 CC for example obesity, dyslipidemia and diabetes.

XX Sequence 1300 AA;

Query Match 29.9%; Score 49; DB 21; Length 1300;
 Best Local Similarity 32.3%; Pred. No. 3.8e+02;
 Matches 10; Conservative 2; Mismatches 7; Indels 12; Gaps 1;

PR	16-AUG-1999;	99US-0149368.	ID	AAG28164 standard; Protein; 67 AA.
PR	17-AUG-1999;	99US-0149175.	XX	
PR	18-AUG-1999;	99US-0149426.	AC	AAG28164;
PR	20-AUG-1999;	99US-0149722.	XX	
PR	20-AUG-1999;	99US-0149723.	DT	17-OCT-2000 (first entry)
PR	20-AUG-1999;	99US-0149929.	XX	
PR	23-AUG-1999;	99US-0149902.	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 33281.
PR	23-AUG-1999;	99US-0149930.	XX	
PR	25-AUG-1999;	99US-0150566.	KW	protein identification; signal transduction pathway; metabolic pathway;
PR	26-AUG-1999;	99US-0150884.	KW	hybridisation assay; genetic mapping; gene expression control; promoter;
PR	27-AUG-1999;	99US-0151065.	KW	termination sequence.
PR	27-AUG-1999;	99US-0151066.	XX	
PR	27-AUG-1999;	99US-0151080.	OS	Arabidopsis thaliana.
PR	30-AUG-1999;	99US-0151303.	XX	
PR	31-AUG-1999;	99US-0151438.	PN	EP1033405-A2.
PR	01-SEP-1999;	99US-0151930.	XX	
PR	07-SEP-1999;	99US-0152363.	PD	06-SEP-2000.
PR	10-SEP-1999;	99US-0153070.	XX	
PR	13-SEP-1999;	99US-0153758.	XX	
PR	15-SEP-1999;	99US-0154018.	PF	25-FEB-2000; 2000EP-0301439.
PR	16-SEP-1999;	99US-0154039.	XX	
PR	20-SEP-1999;	99US-0154779.	PR	25-FEB-1999; 99US-0121825.
PR	22-SEP-1999;	99US-0155139.	PR	05-MAR-1999; 99US-0123180.
PR	23-SEP-1999;	99US-0155486.	PR	09-MAR-1999; 99US-0123548.
PR	24-SEP-1999;	99US-0155659.	PR	23-MAR-1999; 99US-0125788.
PR	28-SEP-1999;	99US-0156458.	PR	25-MAR-1999; 99US-0126264.
PR	29-SEP-1999;	99US-0156596.	PR	29-MAR-1999; 99US-0126785.
PR	04-OCT-1999;	99US-0157117.	PR	01-APR-1999; 99US-0127462.
PR	05-OCT-1999;	99US-0157753.	PR	06-APR-1999; 99US-0128234.
PR	06-OCT-1999;	99US-0157865.	PR	08-APR-1999; 99US-0128714.
PR	07-OCT-1999;	99US-0158029.	PR	16-APR-1999; 99US-0129845.
PR	08-OCT-1999;	99US-0158232.	PR	19-APR-1999; 99US-0130077.
PR	12-OCT-1999;	99US-0158369.	PR	21-APR-1999; 99US-0130449.
PR	13-OCT-1999;	99US-0159293.	PR	23-APR-1999; 99US-0130510.
PR	13-OCT-1999;	99US-0159294.	PR	23-APR-1999; 99US-0130891.
PR	13-OCT-1999;	99US-0159295.	PR	28-APR-1999; 99US-0131449.
PR	14-OCT-1999;	99US-0159329.	PR	30-APR-1999; 99US-0132048.
PR	14-OCT-1999;	99US-0159330.	PR	30-APR-1999; 99US-0132407.
PR	14-OCT-1999;	99US-0159331.	PR	04-MAY-1999; 99US-0132484.
PR	14-OCT-1999;	99US-0159637.	PR	05-MAY-1999; 99US-0132485.
PR	14-OCT-1999;	99US-0159638.	PR	06-MAY-1999; 99US-0132486.
PR	18-OCT-1999;	99US-0159584.	PR	06-MAY-1999; 99US-0132487.
PR	21-OCT-1999;	99US-0160741.	PR	07-MAY-1999; 99US-0132863.
PR	21-OCT-1999;	99US-0160767.	PR	11-MAY-1999; 99US-0134256.
PR	21-OCT-1999;	99US-0160768.	PR	14-MAY-1999; 99US-0134218.
PR	21-OCT-1999;	99US-0160770.	PR	14-MAY-1999; 99US-0134219.
PR	21-OCT-1999;	99US-0160814.	PR	14-MAY-1999; 99US-0134221.
PR	21-OCT-1999;	99US-0160815.	PR	14-MAY-1999; 99US-0134370.
PR	22-OCT-1999;	99US-0160980.	PR	18-MAY-1999; 99US-0134768.
PR	22-OCT-1999;	99US-0160981.	PR	19-MAY-1999; 99US-0134941.
PR	22-OCT-1999;	99US-0160989.	PR	20-MAY-1999; 99US-0135124.
PR	25-OCT-1999;	99US-0161404.	PR	20-MAY-1999; 99US-0135353.
PR	25-OCT-1999;	99US-0161405.	PR	24-MAY-1999; 99US-0135629.
PR	25-OCT-1999;	99US-0161406.	PR	25-MAY-1999; 99US-0136021.
PR	26-OCT-1999;	99US-0161359.	PR	27-MAY-1999; 99US-0136392.
PR	26-OCT-1999;	99US-0161360.	PR	28-MAY-1999; 99US-0136782.
PR	26-OCT-1999;	99US-0161361.	PR	01-JUN-1999; 99US-0137222.
PR	28-OCT-1999;	99US-0161920.	PR	03-JUN-1999; 99US-0137528.
PR	28-OCT-1999;	99US-0161992.	PR	04-JUN-1999; 99US-0137502.
PR	28-OCT-1999;	99US-0161993.	PR	07-JUN-1999; 99US-0137724.
PR	29-OCT-1999;	99US-0162142.	PR	08-JUN-1999; 99US-0138094.
PR			PR	10-JUN-1999; 99US-0138540.
PR			PR	10-JUN-1999; 99US-0138847.
PR			PR	14-JUN-1999; 99US-0139119.
PR			PR	16-JUN-1999; 99US-0139452.
PR			PR	16-JUN-1999; 99US-0139453.
PR			PR	17-JUN-1999; 99US-0139492.
PR			PR	17-JUN-1999; 99US-0139454.
PR			PR	18-JUN-1999; 99US-0139455.
PR			PR	18-JUN-1999; 99US-0139456.
PR			PR	18-JUN-1999; 99US-0139457.
PR			PR	18-JUN-1999; 99US-0139458.
PR			PR	18-JUN-1999; 99US-0139459.
PR			PR	18-JUN-1999; 99US-0139460.

Query Match 29.6%; Score 48.5; DB 21; Length 59;

Best Local Similarity 37.0%; Pred. No. 24;

Matches 10; Conservative 2; Mismatches 10; Indels 5; Gaps 1;

QY 1 KRCINQLCKLPTDSELAPRSXCCHCR 27

Db : | : | | | | | | | | | |
33 rrvpcscvccap-----aprgaccpcor 54

RESULT 29
AAG28164

```
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
```

```
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 29.6%; Score 48.5; DB 21; Length 67;
Best Local Similarity 37.0%; Pred. No. 27;
Matches 10; Conservative 2; Mismatches 10; Indels 5; Gaps 1;

Qy 1 KRCINOLLCKLPTDSELAPEKSCCHCR 27
Db 41 rrvcpcvccap-----aprgaccpcr 62

RESULT 30
AAY23231
ID AAY23231 standard; Protein; 132 AA.
XX
AC AAY23231;
XX
DT 26-AUG-1999 (first entry)
XX
DE Erythrovirus V9 11 kDa protein.

XX Erythrovirus V9; differential diagnosis; parvovirus; infection;
 KW erythrovirus screening; typing; immunoassay.
 XX Erythrovirus.
 XX FR2771751-A1.
 XX 04-JUN-1999.
 XX 03-DEC-1997; 97FR-0015197.
 XX 03-DEC-1997; 97FR-0015197.
 XX (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.
 XX Auguste V, Garbarg CA, Nguyen QT;
 XX WPI; 1999-349543/30.
 DR N-PSDB; AAX81587.
 XX Erythrovirus V9 and its nucleic acid sequences - can be used in the
 PT diagnosis of its infections
 XX Claim 19; Page 59-60; 80pp; French.
 XX The present sequence represents an erythrovirus V9 protein.
 CC Probes and primers derived from erythrovirus V9 polynucleotide
 CC sequences (AAX81580) can be used for differential diagnosis of
 CC erythrovirus (parvovirus) infections by a combination of
 CC amplification and hybridisation assay. The probes can also be
 CC used to assess susceptibility to erythrovirus infection and
 CC for erythrovirus screening and typing. The antibodies can be
 CC used in immunoassays for diagnosis of erythrovirus V9 infections.
 XX Sequence 132 AA;
 SQ
 Query Match 29.6%; Score 48.5; DB 20; Length 132;
 Best Local Similarity 25.0%; Pred. No. 52;
 Matches 10; Conservative 7; Mismatches 8; Indels 15; Gaps 2;
 QY 4 INQLLCKLP-----TSEL-----APRSXCHCRH 28
 DB 27 iyhmycmtpqlgmqnsnttdtdmkslnkncplpkavctchckh 66
 RESULT 31
 AAP93285
 ID AAP93285 standard; protein; 3080 AA.
 XX AAP93285;
 XX 17-DEC-2001 (updated)
 DT 06-APR-1990 (first entry)
 XX Sequence of clone HIV-2 SBL/ISY.
 XX HIV-2; proviral clone HIV-2 SBL/ISY;
 KW Human immunodeficiency virus 2.
 OS USN7331212-N.
 XX 29-AUG-1989.
 XX 31-MAR-1989; 89US-0331212.
 XX 31-MAR-1989; 89US-0331212.
 XX (USSR) US DEPT. HEALTH AND HUMAN SERVICES.
 XX Franchini G, Wong-Staal F, Gallo R;
 PI

XX WPI; 1989-339698/46.
 DR N-PSDB; AAN92119.
 XX Complete human immunodeficiency type 2 proviral clone - used to generate
 PT animal model for function studies of HIV genes in vivo.
 XX Disclosure; Fig. 5; 43pp; English.
 XX The protein is encoded by the third reading frame of HIV-2 SBL/ISY, a
 CC proviral clone of HIV-2.
 CC (Note: Revised entry submitted to correct the patent number format of
 CC US Government-owned NTIS applications to prevent clashes with ongoing US
 CC granted patent numbers. For further information please visit the Derwent
 CC web site at www.derwent.com/dwpi/updates/ntis_us.html.)
 XX Sequence 3080 AA;
 SQ
 Query Match 29.6%; Score 48.5; DB 10; Length 3080;
 Best Local Similarity 47.8%; Pred. No. 9.9e+02;
 Matches 11; Conservative 2; Mismatches 7; Indels 3; Gaps 1;
 QY 3 C1NQILLKLPDSELA PRSXCH 25
 DB 45 cikdliliklprvrskl---svcsch 64
 RESULT 32
 AAP81771
 ID AAP81771 standard; protein; 3210 AA.
 XX AAP81771;
 XX 07-NOV-1990 (first entry)
 DT Deduced sequence encoded by bottom reading frame of cDNA clone
 DE HIV-2 SBL/ISY of HIV related retrovirus strain.
 XX HIV vaccine; HIV strain SBL-6669-85.
 KW Human immunodeficiency virus.
 OS WO8808449-A.
 PN 03-NOV-1988.
 PD 28-APR-1988; 88WO-SE00218.
 PF 28-APR-1987; 87SE-0001765.
 PR (SBL-) SBL STATENS BEKTERI.
 XX Albert J, Biberfeld G, Fenyo EM, Norrby E;
 PI WPI; 1988-322769/45.
 DR N-PSDB; AAN80890.
 XX HIV related human retro-virus strain -
 PT used for obtaining antigens for assays and vaccines and for
 PT prodn. of antibodies for assays
 XX Claim 9; Fig 4; 28pp; English.
 XX Synthetically produced proteins and peptides, characterised in that the
 CC AA sequence is derived from the primary nucleotide sequence of
 CC HIV-2 SBL/ISY or a part thereof, or a degenerate thereof are claimed.
 CC HIV-2 SBL/ISY represents the complete genome of the virus SBL-6699
 CC (=SBL-6689-85). The proviral DNA was obt'd. from a genomic library
 CC constructed from DNA of HUT-78 cells infected with SBL-6669-85 using
 CC the lambda-phage vector EMBL-3. SBL-6669-85 was isolated from lymphocytes
 CC of a West African woman. Protection is requested for the entire genome
 CC disclosed in n80890 and for parts thereof, and corresp. to various genes

CC such as the gag gene (corresp. to nucleotides 547 to 2106), the pol gene
 CC (nucleotides 1827-4931) and the env gene (nucleotides 6144 to 8682), the
 CC corresp. AA sequences and parts thereof and various products derived
 CC therefrom, or use thereof, such as clones prepd. by recombinant vector
 CC method, HIV test devices and methods. X corresponds to the translation of
 CC a stop codon.

SQ Sequence 3210 AA;

Query Match 29.6%; Score 48.5; DB 9; Length 3210;

Best Local Similarity 47.8%; Pred. NO. 1e+03; 7; Indels 3; Gaps 1;

Matches 11; Conservative 2; Mismatches 7; Indels 3; Gaps 1;

OY 3 C1NQLCKLPTDSELA PRSXCH 25

DB 45 clklldllklprskl---svcs h 64

RESULT 33

AAAY77147

ID AAY77147 standard; Protein; 129 AA.

AC AAY77147;

XX AAY77147;

XX 08-MAY-2000 (first entry)

DE Pintada fucata pearl layer substrate protein clone #2.

XX Pearl oyster; pearl layer substrate protein; aragonite crystal;

KW crystallisation; synthetic pearl; production.

XX Pintada fucata.

OS JP3004009-B1.

XX 31-JAN-2000.

XX 26-OCT-1998; 98JP-0304179.

XX 26-OCT-1998; 98JP-0304179.

XX (HAYA/) HAYASHI N.

PA (MASU/) MASUO M.

XX WPI; 2000-149882/14.

DR N-PSDB; AAZ87236.

XX Pearl layer substrate protein for artificial synthesis of pearl - has

PT specific molecular weight of N-terminal and amino acid sequence

XX Claim 2; Page 8-9; 16pp; Japanese.

XX The invention relates to the pearl layer substrate protein (sequences

CC obtained given in AAY77146-Y77152) of the pearl oyster Pintada fucata,

CC and nucleotide sequences encoding this substrate protein

CC (AAZ87235-287241). Also claimed is an N-terminal fragment of the

CC substrate protein isolated from the pearl layer (AAY77145). This peptide

CC fragment was used as the basis for the design of degenerate

CC oligonucleotide probes (AAZ87223-287224) for the detection of nucleotides

CC encoding portions of the substrate protein in a Pintada fucata mantle

CC cDNA library. The clones isolated were cloned into a vector

CC (pGEM), amplified using T7 promoter and SP6 promoter PCR primers

CC (AAZ87225-287226) and sequenced. The sequences of two partial cDNAs

CC obtained are given in AAZ87242 and AAZ87243. These were then extended via

CC the RACE (rapid amplification of cDNA ends) procedure using RACE primers

CC PCR primers AAZ87229-287232 to give the full-length cDNA clones. The

CC pearl layer substrate protein acts as a nucleus for aragonite crystal

CC formation during the formation of a pearl. Use of the pearl layer

CC substrate protein may be used in the production of synthetic pearls.

CC Currently, large-scale pearl production is achieved via cultivation of

CC pearl oysters; however, pearl formation can take up to two years.

CC Sequences AAY77146-Y77152 represent Pintada fucata pearl layer substrate
 CC protein sequences encoded by full-length cDNA clones #1-#7 (AAZ87235-
 CC AAZ87341).

SQ Sequence 129 AA;

Query Match 29.3%; Score 48; DB 21; Length 129;

Best Local Similarity 34.4%; Pred. No. 59;

Matches 11; Conservative 3; Mismatches 14; Indels 4; Gaps 1;

OY 1 KRCINQLCKLPTDSELA PR----SXCHCRH 28

DB 27 kkcgrysycwipydierrdrngkckccfcrn 58

RESULT 34

AAE03277

ID AAE03277 standard; Protein; 172 AA.

XX AAE03277;

XX 10-AUG-2001 (first entry)

DE Human gene 13 encoded secreted protein fragment, SEQ ID NO:130.

XX Human; secreted protein; proliferative disorder; cancer; tumour; asthma;

KW foetal abnormality; developmental abnormality; haematopoietic disorder;

KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;

KW Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;

KW psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;

KW inflammation; neurological disorder; Alzheimer's disease; food additive;

KW angiotensin-related disorder; kidney disorder; gastrointestinal disorder; allergy;

KW pregnancy-related disorder; endocrine disorder; infection; wound healing;

KW cell culture; chemotaxis; vulnery; binding partner identification;

KW gene therapy.

XX Homo sapiens.

OS WO200134644-A1.

PN 17-MAY-2001.

PD 08-NOV-2000; 2000WO-US30679.

XX 12-NOV-1999; 99US-0164834.

PR 04-AUG-2000; 2000US-0224007.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Ruben SM, Komatsoulis GA, Olsen HS, Duan RD, Ebner R;

PI WPI; 2001-329070/34.

XX Isolated nucleic acid molecule encoding a human secreted protein is

PT used in preventing, treating or ameliorating a medical condition -

XX Disclosure; Page 47; 499pp; English.

XX AAD07655-AAD07695 represent cDNAs corresponding to 15 human secreted

CC protein genes, and AAE03202-AAE03242 represent the proteins they encode.

CC AAE03243-AAE03280 represent human secreted protein fragments or variants.

CC The secreted proteins and their genes are useful for preventing, treating

CC or ameliorating medical conditions, e.g., by protein or gene therapy.

CC Pathological conditions can be diagnosed by determining the amount of the

CC new protein in a sample or by determining the presence of mutations in

CC the new genes. Specific uses are described for each of the 15 genes,

CC based on the tissues in which they are most highly expressed, and include

CC developing products for the diagnosis or treatment of proliferative

CC disorders, cancer, tumours, foetal and developmental abnormalities,

CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune

CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,

CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),

CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
CC angioenic disorders, kidney disorders, gastrointestinal disorders,
CC pregnancy-related disorders, endocrine disorders, and infections. The
CC proteins can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, for supporting cell culture of primary tissues,
CC to regenerate tissues, to identify their cognate ligands or binding
CC partners, and in chemotaxis, and can be used as a food additive or
CC preservative to modify storage properties. Antibodies specific for a
CC protein of the invention can be used in alleviating symptoms associated
CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,
CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA).
CC The present sequence represents a human secreted protein fragment
CC referred to in the disclosure of the invention.

XX Sequence 172 AA;

Query Match 29.3%; Score 48; DB 22; Length 172;
Best Local Similarity 75.0%; Pred. No. 77;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 19 PRSXCCHC 26
||: ||||
Db 2 prascchc 9

RESULT 35

AAAG54677
ID AAG54677 standard; Protein; 296 AA.

XX AC AAG54677;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 69755.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 21-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.

23-JUL-1999; 99US-0145224.
 26-JUL-1999; 99US-0145276.
 27-JUL-1999; 99US-0145913.
 27-JUL-1999; 99US-0145918.
 27-JUL-1999; 99US-0145919.
 28-JUL-1999; 99US-0145951.
 02-AUG-1999; 99US-0146386.
 02-AUG-1999; 99US-0146388.
 03-AUG-1999; 99US-0146389.
 03-AUG-1999; 99US-0147038.
 04-AUG-1999; 99US-0147204.
 04-AUG-1999; 99US-0147302.
 05-AUG-1999; 99US-0147192.
 05-AUG-1999; 99US-0147260.
 06-AUG-1999; 99US-0147303.
 06-AUG-1999; 99US-0147416.
 09-AUG-1999; 99US-0147493.
 09-AUG-1999; 99US-0147935.
 10-AUG-1999; 99US-0148171.
 11-AUG-1999; 99US-0148319.
 12-AUG-1999; 99US-0148341.
 13-AUG-1999; 99US-0148565.
 13-AUG-1999; 99US-0148684.
 16-AUG-1999; 99US-0149368.
 17-AUG-1999; 99US-0149175.
 18-AUG-1999; 99US-0149426.
 20-AUG-1999; 99US-0149722.
 20-AUG-1999; 99US-0149723.
 20-AUG-1999; 99US-0149929.
 23-AUG-1999; 99US-0149902.
 23-AUG-1999; 99US-0149930.
 25-AUG-1999; 99US-0150566.
 26-AUG-1999; 99US-0150884.
 27-AUG-1999; 99US-0151065.
 27-AUG-1999; 99US-0151066.
 27-AUG-1999; 99US-0151080.
 30-AUG-1999; 99US-0151303.
 31-AUG-1999; 99US-0151438.
 01-SEP-1999; 99US-0151930.
 07-SEP-1999; 99US-0152363.
 10-SEP-1999; 99US-0153070.
 13-SEP-1999; 99US-0153758.
 15-SEP-1999; 99US-0154018.
 16-SEP-1999; 99US-0154039.
 20-SEP-1999; 99US-0154779.
 22-SEP-1999; 99US-0155139.
 23-SEP-1999; 99US-0155486.
 24-SEP-1999; 99US-0155659.
 28-SEP-1999; 99US-0156458.
 29-SEP-1999; 99US-0156596.
 04-OCT-1999; 99US-0157117.
 05-OCT-1999; 99US-0157753.
 06-OCT-1999; 99US-0157865.
 07-OCT-1999; 99US-0158029.
 08-OCT-1999; 99US-0158232.
 12-OCT-1999; 99US-0158369.
 13-OCT-1999; 99US-0159293.
 13-OCT-1999; 99US-0159294.
 13-OCT-1999; 99US-0159295.
 14-OCT-1999; 99US-0159329.
 14-OCT-1999; 99US-0159330.
 14-OCT-1999; 99US-0159331.
 14-OCT-1999; 99US-0159637.
 14-OCT-1999; 99US-0159638.
 18-OCT-1999; 99US-0159584.
 21-OCT-1999; 99US-0160741.
 21-OCT-1999; 99US-0160767.
 21-OCT-1999; 99US-0160768.
 21-OCT-1999; 99US-0160770.
 21-OCT-1999; 99US-0160814.
 21-OCT-1999; 99US-0160815.
 22-OCT-1999; 99US-0160980.
 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.
 Query Match 29.3%; Score 48; DB 21; Length 296;
 Best Local Similarity 57.9%; Pred. No. 1.3e+02;
 Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
 QY 3 CINOLCKLPTDSELAPRS 21
 Db 196 cinksiakiplsvedaars 214
 |||: |||| | ||
 RESULT 36
 AAB36836
 ID AAB36836 standard; Protein; 1268 AA.
 XX
 AC AAB36836;
 XX
 DT 20-FEB-2001 (first entry)
 XX
 DE Human insulin receptor-related receptor protein.
 XX
 KW Insulin receptor-related receptor; IRRR; chromosome 1q21-q24;
 KW obesity; dyslipidemia; diabetes.
 XX
 OS Homo sapiens.
 XX
 PN WO200065090-A2.
 XX
 PD 02-NOV-2000.
 XX
 PF 19-APR-2000; 2000WO-US10644.
 XX
 PR 22-APR-1999; 99US-0296906.
 PR 22-JUN-1999; 99US-0337976.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 XX Lok S, Whitmore TE;
 XX
 DR WPI; 2000-687365/67.
 DR N-PSDB; AAC68339.
 XX
 XX Detecting a chromosome 1q21-q24 abnormality for diagnosing metabolic
 disease, such as human obesity and diabetic disorders, comprises
 examining insulin receptor-related receptor gene and its gene products
 .
 XX
 XX Disclosure; Page 69-73; 111pp; English.
 PS
 XX The present invention relates to insulin receptor-related receptor
 CC (IRRR). Mutations in this gene indicate a chromosome 1q21-q24
 CC abnormality. IRRR polypeptides and DNA may be useful in the diagnosis of
 CC of disorders associated with abnormal expression of the IRRR protein,
 CC for example obesity, dyslipidemia and diabetes.
 XX
 SQ Sequence 1268 AA;

Query Match 29.3%; Score 48; DB 21; Length 1268;
 Best Local Similarity 26.2%; Pred. No. 5e+02;
 Matches 11; Conservative 1; Mismatches 12; Indels 18; Gaps 1;
 QY 6 QLLCKLPTDSELAPRSXCCH-----CRHL 29

Db 182 qrvpcphgmactargechtcclgscsqedpracvacrhl 223
 RESULT 37
 ID ABB11702 standard; peptide; 1268 AA.
 AC ABB11702;
 XX
 DT 11-JAN-2002 (first entry)
 XX
 DE Human insulin receptor-related protein homologue, SEQ ID NO:2072.
 KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytostatic; osteopathic; vasotrophic; cardiant; virucide; antibacterial;
 KW antifungal; vulnery; antiulcer.
 OS Homo sapiens.
 XX
 XX WO200157188-A2.
 PN
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US03800.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Drmanac RT;
 PI
 XX WPI; 2001-457740/49.
 DR N-PSDB; ABA08946.
 XX
 XX Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -
 XX
 PS Claim 20; Page 229-230; 1963pp; English.
 XX
 CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell

CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention.
 XX
 SQ Sequence 1268 AA;
 Query Match 29.3%; Score 48; DB 22; Length 1268;
 Best Local Similarity 26.2%; Pred No. 5e+02;
 Matches 11; Conservative 1; Mismatches 12; Indels 18; Gaps 1;
 QY 6 QLCLKLPDSELA PRSXCH-----CRHL 29
 | : | | |
 Db 182 qrvpcphgmactargechtcclgscsqedpracvacrhl 223
 RESULT 38
 AAB36840
 ID AAB36840 standard; Protein; 1297 AA.
 XX
 AC AAB36840;
 XX
 DT 20-FEB-2001 (first entry)
 XX
 DE Human insulin receptor-related receptor protein with signal peptide.
 KW Insulin receptor-related receptor; IRRR; chromosome 1q21-q24;
 KW obesity; dyslipidemia; diabetes.
 XX
 OS Homo sapiens.
 XX
 PN WO200065090-A2.
 XX
 PD 02-NOV-2000.
 XX
 PF 19-APR-2000; 2000WO-US10644.
 XX
 PR 22-APR-1999; 99US-0296906.
 PR 22-JUN-1999; 99US-0337976.
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 XX Lok S, Whitmore TE;
 PI
 XX WPI; 2000-687365/67.
 XX
 PT Detecting a chromosome 1q21-q24 abnormality for diagnosing metabolic
 PT disease, such as human obesity and diabetic disorders, comprises
 PT examining insulin receptor-related receptor gene and its gene products
 PT -
 XX
 PS Claim 4; Page 97-101; 111pp; English.
 XX
 CC The present invention relates to insulin receptor-related receptor
 CC (IRRR). Mutations in this gene indicate a chromosome 1q21-q24
 CC abnormality. IRRR polypeptides and DNA may be useful in the diagnosis of
 CC disorders associated with abnormal expression of the IRRR protein,
 CC for example obesity, dyslipidemia and diabetes.
 XX
 SQ Sequence 1297 AA;

Query Match 29.3%; Score 48; DB 21; Length 1297;
Best Local Similarity 26.2%; Pred. No. 5.1e+02;
Matches 11; Conservative 1; Mismatches 12; Indels 18; Gaps 1;

QY 6 QLLCKLPDSELA PRSXCH-----CRHL 29
| : | | | |
Db 211 qrvcpghmactargechtclgscsqedpracvacrhl 252

RESULT 39
AAU58496
ID AAU58496 standard; Protein; 70 AA.

XX AC AAU58496;

XX DT 27-FEB-2002 (first entry)

XX DE Propionibacterium acnes immunogenic protein #19392.

XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypervitosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

XX OS Propionibacterium acnes.

XX PN WO200181581-A2.

XX PD 01-NOV-2001.

XX PF 20-APR-2001; 2001WO-US12865.

XX PR 21-APR-2000; 2000US-199047P.

XX PR 02-JUN-2000; 2000US-208841P.

XX PR 07-JUL-2000; 2000US-216747P.

XX PA (CORI-) CORIXA CORP.

XX PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX PA WPI; 2001-616774/71.

XX DR N-PSDB; AAS59591.

XX PT Propionibacterium acnes polypeptides and nucleic acids useful for

XX PT vaccinating against and diagnosing infections, especially useful for

XX PT treating acne vulgaris.

XX PS Example 1; SEQ ID No 19691; 1069pp; English.

XX CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX CC polypeptides. The proteins and their associated DNA sequences are used in
XX CC the treatment, prevention and diagnosis of medical conditions caused by
XX CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX CC pustulosis, hypervitosis and osteomyelitis), uveitis and endophthalmitis.
XX CC P. acnes is also involved in infections of bone, joints and the central
XX CC nervous system, however it is particularly involved in the inflammatory
XX CC lesions associated with acne vulgaris. A method for detecting the
XX CC presence or absence of P. acnes in a patient comprises contacting a
XX CC sample with a binding agent that binds to the proteins of the invention
XX CC and determining the amount of bound protein in the sample. The
XX CC polypeptides may be used as antigens in the production of antibodies
XX CC specific for P. acnes proteins. These antibodies can be used to
XX CC downregulate expression and activity of P. acnes polypeptides and
XX CC therefore treat P. acnes infections. The antibodies may also be used as
XX CC diagnostic agents for determining P. acnes presence, for example, by
XX CC enzyme linked immunosorbent assay (ELISA).

XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 70 AA;

Query Match 29.0%; Score 47.5; DB 22; Length 70;
Best Local Similarity 44.0%; Pred. No. 38;
Matches 11; Conservative 3; Mismatches 8; Indels 3; Gaps 1;

QY 3 CINQLCKLP---TDSLEAPRSXCC 24
| : | | | | | | | | | |
Db 16 clsqslktpssltnsasgrsvsc 40

RESULT 40

AAU26136

ID AAY26136 standard; Protein; 182 AA.

XX AC AAY26136;

XX DT 22-OCT-1999 (first entry)

XX DE Phospholipase A2 inhibitor-NAI alpha ii.

XX KW Phospholipase A2 inhibitor; PLA2 inhibitor; NAI alpha ii; alpha chain;
KW anti-inflammatory activity; recombinant PLA2 inhibitor; prophylactic;
KW therapeutic application; rheumatoid arthritis; osteoarthritis; asthma;
KW allergic reaction; psoriasis; multiple organ failure; acute pancreatitis;
KW acute lung failure; septic shock; adult respiratory distress syndrome;
KW toxin.

XX OS Notechis ater.

XX PN WO9929726-A1.

XX PD 17-JUN-1999.

XX PF 30-NOV-1998; 98WO-AU00992.

XX PR 05-DEC-1997; 97AU-0000767.

XX PA (ACTI-) ACTIVE PLA R & D PTY LTD.

XX PA (APEL-) APELDA PLA R & D PTY LTD.

XX PA (EDZE-) EDZELL PLA R & D PTY LTD.

XX PA (FLAI-) FLAIR PLA R & D PTY LTD.

XX PA (HERA-) HERACLES PLA R & D PTY LTD.

XX PA (HSCP-) HSC PLA PTY LTD.

XX PA (NORT-) NORTHMOOR PLA R & D PTY LTD.

XX PI Broady KW, Hains PG;

XX DR WPI; 1999-385577/32.

XX DR N-PSDB; AAX80576.

XX PT Broad-spectrum phospholipase A2 inhibitor derived from snake venom

XX PS Claim 10; Pages 93-94; 161pp; English.

XX CC The present sequence is an isolated broad-spectrum phospholipase A2
XX CC (PLA2) inhibitor, NAI alpha ii. It comprises alpha-chain from isoform ii
XX CC of PLA2 inhibitor isolated from Notechis ater. It is capable of
XX CC inhibiting two or more of PLA2 Type I, II and/or III enzymes. The PLA2
XX CC inhibitor has anti-inflammatory activity. The isolated PLA2 inhibitor or
XX CC recombinant PLA2 inhibitor are useful in a wide range of prophylactic and
XX CC therapeutic applications. The phospholipase inhibitors are useful for the
XX CC treatment of the phospholipase-related symptom(s) of rheumatoid
XX CC arthritis, osteoarthritis, asthma, allergic reaction, psoriasis, multiple
XX CC organ failure, acute pancreatitis, acute lung failure, septic shock,
XX CC adult respiratory distress syndrome or the toxic effects of toxins (e.g.
XX CC due to insect bites) in humans or animals.

XX SQ Sequence 182 AA;

Query Match 29.0%; Score 47.5; DB 20; Length 182;

Best Local Similarity 35.5%; Pred. NO. 94;
Matches 11; Conservative 3; Mismatches 10; Indels 7; Gaps 1;
QY 1 KRCINOLICKLPT-----DSELAPRSXCC 24
Db 47 rKcfsscklerfdinighdsylgrthcc 77

Search completed: August 22, 2002, 07:50:21
Job time: 184 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 07:47:42 ; Search time 21.93 seconds
(without alignments)
32.300 Million cell updates/sec

Title: US-09-856-199-1

Perfect score: 164

Sequence: 1 KRCINQLLCKLPTDSELRSCXCHCRHL 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued_Patents_AA.*

1: /cgn2.6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2.6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2.6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2.6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2.6/ptodata/2/iaa/PTCUS_COMB.pep.*

6: /cgn2.6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	31.7	430	3	US-08-997-897-2
2	52	31.7	430	4	US-09-156-836B-2
3	48.5	29.6	3080	6	5223423-4
4	47	28.7	659	4	US-09-268-347-46
5	46	28.0	158	4	US-08-679-493A-24
6	46	28.0	507	1	US-08-484-493-12
7	46	28.0	507	1	US-08-484-494-12
8	46	28.0	507	2	US-08-345-212-12
9	46	28.0	507	4	US-09-249-003-8
10	46	28.0	509	1	US-08-445-586-8
11	45	27.4	82	1	US-07-965-674-1
12	45	27.4	82	5	PCT-US93-09523-1
13	45	27.4	1554	2	US-08-705-625-3
14	45	27.4	1554	3	US-09-010-998-6
15	45	27.4	1554	4	US-09-220-574-3
16	44	26.8	1912	4	US-08-913-832A-2
17	43	26.2	78	3	US-08-946-026-26
18	43	26.2	415	2	US-08-815-718-2
19	43	26.2	540	2	US-08-884-072-3
20	43	26.2	540	4	US-09-212-168-3
21	43	26.2	588	3	US-08-946-026-3
22	43	26.2	1059	4	US-09-134-513-2
23	42	25.6	21	2	US-08-637-759B-485
24	42	25.6	21	3	US-08-871-355A-485
25	42	25.6	21	4	US-09-201-945-485
26	42	25.6	62	1	US-07-965-674-9
27	42	25.6	62	5	PCT-US93-09523-9

28	42	25.6	70	6	5318899-16
29	42	25.6	351	4	US-09-524-162-2
30	42	25.6	351	5	PCT-US91-06418-1
31	42	25.6	445	2	US-08-985-090-2
32	42	25.6	445	3	US-09-165-543-2
33	42	25.6	445	4	US-09-167-354-7
34	42	25.6	473	2	US-08-759-581B-15
35	42	25.6	473	4	US-09-304-711-15
36	42	25.6	499	4	US-09-049-672A-1
37	42	25.6	751	2	US-08-836-443-3
38	42	25.6	855	2	US-09-027-337-2
39	41.5	25.3	688	4	US-09-113-750A-3
40	41.5	25.3	814	4	US-09-813-819-4
41	41.5	25.3	814	4	US-09-920-048-4
42	41.5	25.3	855	4	US-09-813-819-2
43	41.5	25.3	855	4	US-09-920-048-2
44	41.5	25.3	1835	4	US-09-404-650-5
45	41	25.0	72	6	5318899-14
46	41	25.0	83	1	US-07-965-674-2
47	41	25.0	83	1	US-07-965-674-4
48	41	25.0	83	5	PCT-US93-09523-2
49	41	25.0	83	5	PCT-US93-09523-4
50	41	25.0	83	6	5318899-20
51	41	25.0	84	1	US-07-965-674-3
52	41	25.0	84	5	PCT-US93-09523-3
53	41	25.0	97	3	US-08-946-026-29
54	41	25.0	457	1	US-08-264-101-4
55	41	25.0	457	2	US-08-765-243-4
56	41	25.0	457	5	PCT-US95-07295-4
57	41	25.0	735	2	US-08-765-243-6
58	41	25.0	735	5	PCT-US95-07295-6
59	41	25.0	798	2	US-07-728-215-30
60	41	25.0	798	4	US-08-938-085A-30
61	41	25.0	1463	1	US-08-220-603A-11
62	41	25.0	1835	3	US-08-836-325-15
63	41	25.0	3077	6	5223423-2
64	40.5	24.7	108	1	US-08-485-359-2
65	40.5	24.7	108	1	US-08-569-594-2
66	40.5	24.7	108	5	PCT-US96-08815-2
67	40.5	24.7	181	3	US-08-848-580-12
68	40.5	24.7	227	1	US-08-414-625-4
69	40.5	24.7	381	2	US-09-193-877-2
70	40.5	24.7	488	1	US-08-554-659-2
71	40.5	24.7	488	1	US-08-554-659-4
72	40.5	24.7	2431	1	US-07-920-281C-2
73	40.5	24.7	2431	1	US-08-466-277-2
74	40	24.4	44	2	US-08-557-309B-65
75	40	24.4	71	6	5318899-11
76	40	24.4	71	6	5318899-13
77	40	24.4	72	6	5318899-8
78	40	24.4	72	6	5318899-9
79	40	24.4	73	6	5318899-2
80	40	24.4	73	6	5318899-3
81	40	24.4	73	6	5318899-4
82	40	24.4	73	6	5318899-5
83	40	24.4	73	6	5318899-6
84	40	24.4	73	6	5318899-10
85	40	24.4	74	6	5318899-12
86	40	24.4	95	6	5318899-26
87	40	24.4	96	2	US-08-807-861A-35
88	40	24.4	96	2	US-08-470-868A-35
89	40	24.4	96	3	US-09-210-681-35
90	40	24.4	96	3	US-08-946-719A-35
91	40	24.4	128	2	US-08-557-309B-41
92	40	24.4	128	3	US-08-834-306-41
93	40	24.4	128	4	US-08-993-674A-41
94	40	24.4	560	2	US-08-808-931-18
95	40	24.4	560	3	US-08-808-323-18
96	40	24.4	560	3	US-09-050-603A-18
97	40	24.4	560	3	US-09-102-420B-18
98	40	24.4	560	4	US-09-497-698-18
99	39.5	24.1	700	2	US-08-568-459A-10
100	39.5	24.1	700	2	US-08-487-826B-10


```

US-08-679-493A-24
: Sequence 24, Application US/08679493A
: Patent No. 6303295
: GENERAL INFORMATION:
: APPLICANT: Taylor, Ethan W.
: TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS
: FILE REFERENCE: 55-95
: CURRENT APPLICATION NUMBER: US/08/679.493A

```

; CURRENT FILING DATE: 1996-07-12
; PRIOR APPLICATION NUMBER: 60/001203
; PRIOR FILING DATE: 1995-07-14
; PRIOR APPLICATION NUMBER: 60/003,112
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(158)
; OTHER INFORMATION: X is selenocysteine.
US-08-679-493A-24

Query Match 28.0%; Score 46; DB 4; Length 158;
Best Local Similarity 28.6%; Pred. No. 21;
Matches 6; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 9 CKLPTDSELAPRSXKCHCRHL 29
|::| |::| |::| |::|
Db 111 CRPPSMATCQRTWCCTCSRM 131

RESULT 6
US-08-484-493-12
; Sequence 12, Application US/08484493
; Patent No. 5728381
; GENERAL INFORMATION:
; APPLICANT: Wilson, Peter J
; APPLICANT: Morris, Charles P
; APPLICANT: Anson, Donald S
; APPLICANT: Occhiodoro, Teresa
; APPLICANT: Bielicki, Julie
; APPLICANT: Clements, Peter R
; APPLICANT: Hopwood, John J
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/484,493
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 991,973
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8416Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 507 amino acids
; TYPE: amino acid

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-484-493-12

Query Match 28.0%; Score 46; DB 1; Length 507;
Best Local Similarity 75.0%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 19 PRSXCHC 26
|| ||||
Db 495 PRPACCHC 502

RESULT 7
US-08-484-494-12
; Sequence 12, Application US/08484494
; Patent No. 5798239
; GENERAL INFORMATION:
; APPLICANT: Wilson, Peter J
; APPLICANT: Morris, Charles P
; APPLICANT: Anson, Donald S
; APPLICANT: Occhiodoro, Teresa
; APPLICANT: Bielicki, Julie
; APPLICANT: Clements, Peter R
; APPLICANT: Hopwood, John J
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/484,494
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 991,973
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8416Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 507 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-484-494-12

Query Match 28.0%; Score 46; DB 1; Length 507;
Best Local Similarity 75.0%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 19 PRSXCHC 26
|| ||||

Db 495 PRPACCHC 502

RESULT 8

US-08-345-212-12
; Sequence 12, Application US/08345212
; Patent No. 5932211
; GENERAL INFORMATION:
; APPLICANT: Wilson, Peter J
; APPLICANT: Morris, Charles P
; APPLICANT: Anson, Donald S
; APPLICANT: Occhiodoro, Teresa
; APPLICANT: Bielicki, Julie
; APPLICANT: Clements, Peter R
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
; IDURONATE 2-SULFATASE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/345,212
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 991,973
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8416Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 507 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-345-212-12

Query Match 28.0%; Score 46; DB 2; Length 507;
Best Local Similarity 75.0%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 19 PRSXCCHC 26

Db 495 PRPACCHC 502

RESULT 9

US-09-249-003-12
; Sequence 12, Application US/09249003
; Patent No. 6153188
; GENERAL INFORMATION:
; APPLICANT: Wilson, Peter J
; APPLICANT: Morris, Charles P
; APPLICANT: Anson, Donald S
; APPLICANT: Occhiodoro, Teresa

; APPLICANT: Bielicki, Julie
; APPLICANT: Clements, Peter R
; APPLICANT: Hopwood, John J
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
; IDURONATE 2-SULFATASE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/249,003
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/991,973
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8416Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 507 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-249-003-12

Query Match 28.0%; Score 46; DB 4; Length 507;
Best Local Similarity 75.0%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 19 PRSXCCHC 26

Db 495 PRPACCHC 502

RESULT 10

US-08-445-586-8
; Sequence 8, Application US/08445586
; Patent No. 5627050
; GENERAL INFORMATION:
; APPLICANT: Takeshita, Sunao
; APPLICANT: Ito, Toshimi
; APPLICANT: Otawara-Hamamoto, Yoko
; APPLICANT: Amann, Egon
; TITLE OF INVENTION: Bone-Related Sulfatase-Like Protein and
; Process for its Production
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3314
; COMPUTER READABLE FORM:

;/ MEDIUM TYPE: Floppy disk
;/ COMPUTER: IBM PC compatible
;/ OPERATING SYSTEM: PC-DOS/MS-DOS
;/ SOFTWARE: PatentIn Release #1.0, Version #1.25
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/445,586
;/ FILING DATE:
;/ CLASSIFICATION: 435
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: US 08/111,887
;/ FILING DATE: 26-AUG-1993
;/ APPLICATION NUMBER: JP 230030/92
;/ FILING DATE: 28-AUG-1992
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: JP 324034/92
;/ FILING DATE: 03-DEC-1992
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Forman, David S.
;/ REGISTRATION NUMBER: 33,694
;/ REFERENCE/DOCKET NUMBER: 02481.1322-00000
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: 202-408-4000
;/ TELEFAX: 202-408-4400
;/ INFORMATION FOR SEQ ID NO: 8:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 509 amino acids
;/ TYPE: amino acid
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: peptide
;/ US-08-445-586-8

Query Match 28.0%; Score 46; DB 1; Length 509;
Best Local Similarity 75.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 19 PRSXCCHC 26
DB 497 PRPACCHC 504

RESULT 11
US-07-965-674-1
;/ Sequence 1, Application US/07965674
;/ Patent No. 5380646
;/ GENERAL INFORMATION:
;/ APPLICANT: Knight, Linda C.
;/ TITLE OF INVENTION: Thrombus Detection Using
;/ TITLE OF INVENTION: Radiolabelled Disintegrins
;/ NUMBER OF SEQUENCES: 16
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: Temple University - Of The Common-
;/ ADDRESSEE: wealth System of Higher Education
;/ STREET: 406 University Services Building
;/ CITY: Philadelphia
;/ STATE: Pennsylvania
;/ COUNTRY: U.S.A.
;/ ZIP: 19122
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Diskette, 3.50 inch, 720 KB
;/ COMPUTER: IBM PS/2
;/ OPERATING SYSTEM: MS-DOS
;/ SOFTWARE: Wordperfect 5.1
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/07/965,674
;/ FILING DATE: 19921019
;/ CLASSIFICATION: 514
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER:
;/ FILING DATE:
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Monaco, Daniel A.
;/ REGISTRATION NUMBER: 30,480

;/ REFERENCE/DOCKET NUMBER: 6056-173
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (215) 568-8383
;/ TELEFAX: (215) 568-5549
;/ TELEX: No. 5380646e
;/ INFORMATION FOR SEQ ID NO: 1:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 82 amino acids
;/ TYPE: AMINO ACID
;/ TOPOLOGY: linear
;/ US-07-965-674-1

Query Match 27.4%; Score 45; DB 1; Length 82;
Best Local Similarity 34.6%; Pred. No. 15;
Matches 9; Conservative 2; Mismatches 15; Indels 0; Gaps 0;

QY 2 RCINQLLCKLPDSELA PRSXCHCR 27
DB 26 RCNNAATCKLTPGSQCNHGCCDQCK 51

RESULT 12
PCT-US93-09523-1
;/ Sequence 1, Application PC/TUS9309523
;/ GENERAL INFORMATION:
;/ APPLICANT: Temple University - Of The Commonwealth
;/ APPLICANT: System of Higher Education
;/ APPLICANT: Knight, Linda C. and
;/ APPLICANT: Maurer, Alan H.
;/ TITLE OF INVENTION: Thrombus Detection Using
;/ TITLE OF INVENTION: Radiolabelled Disintegrins
;/ NUMBER OF SEQUENCES: 16
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: Temple University - Of The
;/ ADDRESSEE: Commonwealth System of Higher Education
;/ STREET: 406 University Services Building
;/ CITY: Philadelphia
;/ STATE: Pennsylvania
;/ COUNTRY: U.S.A.
;/ ZIP: 19122
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Diskette, 3.50 inch, 720 KB
;/ COMPUTER: IBM PS/2
;/ OPERATING SYSTEM: MS-DOS
;/ SOFTWARE: Wordperfect 5.1
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: PCT/US93/09523
;/ FILING DATE:
;/ CLASSIFICATION:
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: 965,674
;/ FILING DATE: 19 October 1992
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Monaco, Daniel A.
;/ REGISTRATION NUMBER: 30,480
;/ REFERENCE/DOCKET NUMBER: 6056-173 PC
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (215) 568-8383
;/ TELEFAX: (215) 568-5549
;/ TELEX: None
;/ INFORMATION FOR SEQ ID NO: 1:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 82 amino acids
;/ TYPE: amino acid
;/ TOPOLOGY: linear
;/ PCT-US93-09523-1

Query Match 27.4%; Score 45; DB 5; Length 82;
Best Local Similarity 34.6%; Pred. No. 15;
Matches 9; Conservative 2; Mismatches 15; Indels 0; Gaps 0;

QY 2 RCINQLKLPDSELA PRSXCHCR 27
||| ||| | :
Db 26 RCCNATCKLTPGSCNHCDCDOCK 51

RESULT 13

US-08-705-625-3
; Sequence 3, Application US/08705625
; Patent No. 5908756
; GENERAL INFORMATION:
; APPLICANT: Snyder, Solomon H.
; APPLICANT: Jaffrey, Samie R.
; TITLE OF INVENTION: Protein Inhibitor of Neuronal Nitric
; TITLE OF INVENTION: Oxide Synthase
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff
; STREET: 1001 G Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20001-4597
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/705,625
; FILING DATE: 30-AUG-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.57071
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1554 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Rattus norvegicus
US-08-705-625-3

Query Match 27.4%; Score 45; DB 2; Length 1554;
Best Local Similarity 40.9%; Pred. No. 2.8e+02;
Matches 9; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 6 QLLCKLPDSELA PRSXCHCR 27
||| ||| | :
Db 1451 QTLNLSGTPGPRASCPCR 1472

RESULT 14

US-09-010-998-6
; Sequence 6, Application US/09010998
; Patent No. 6103872
; GENERAL INFORMATION:
; APPLICANT: Snyder, Solomon
; APPLICANT: Jaffrey, Samie
; APPLICANT: Snowman, Adele
; APPLICANT: Eliasson, Mikael
; APPLICANT: Cohen, No. 6103872m
; TITLE OF INVENTION: CAPON, a protein that binds
; TITLE OF INVENTION: neuronal nitric oxide synthase
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff
; STREET: 1001 G Street, NW

Query Match 27.4%; Score 45; DB 3; Length 1554;
Best Local Similarity 40.9%; Pred. No. 2.8e+02;
Matches 9; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 6 QLLCKLPDSELA PRSXCHCR 27
||| ||| | :
Db 1451 QTLNLSGTPGPRASCPCR 1472

CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/010,998
FILING DATE: 22-JAN-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32141
REFERENCE/DOCKET NUMBER: 01107.73424
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1554 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6103872e
US-09-010-998-6

Query Match 27.4%; Score 45; DB 3; Length 1554;
Best Local Similarity 40.9%; Pred. No. 2.8e+02;
Matches 9; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 6 QLLCKLPDSELA PRSXCHCR 27
||| ||| | :
Db 1451 QTLNLSGTPGPRASCPCR 1472

RESULT 15

US-09-220-574-3
; Sequence 3, Application US/09220574
; Patent No. 6168926
; GENERAL INFORMATION:
; APPLICANT: Snyder, Solomon H.
; APPLICANT: Jaffrey, Samie R.
; TITLE OF INVENTION: Protein Inhibitor of Neuronal Nitric
; TITLE OF INVENTION: Oxide Synthase
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff
; STREET: 1001 G Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20001-4597
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/220,574
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/705,625
; FILING DATE: 30-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.

; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.57071
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1554 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Rattus norvegicus
US-09-220-574-3

Query Match 27.4%; Score 45; DB 4; Length 1554;
Best Local Similarity 40.9%; Pred. No. 2.8e+02;
Matches 9; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

Qy 6 QLLCKLPTDSELAPRSXCCHCR 27
Db 1451 QTLNLSGTPCGPRASCPGR 1472

RESULT 16
US-08-913-832A-2
; Sequence 2, Application US/08913832A
; Patent No. 6329517
; GENERAL INFORMATION:
; APPLICANT: Seelig, Hans Peter
; APPLICANT: Renz, Manfred
; TITLE OF INVENTION: DERMATOMYOSITIS-SPECIFIC AUTO-ANTIGEN
; FILE REFERENCE: 8484-0030-999
; CURRENT APPLICATION NUMBER: US/08/913.832A
; CURRENT FILING DATE: 1998-01-12
; PRIOR APPLICATION NUMBER: PCT/DE96/00444
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1912
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-913-832A-2

Query Match 26.8%; Score 44; DB 4; Length 1912;
Best Local Similarity 38.1%; Pred. No. 4.7e+02;
Matches 8; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 3 CINQLCKLPTDSELAPRSXC 23
Db 475 CLNPPPEIPNGEWLCPRCTC 495

RESULT 17
US-08-946-026-26
; Sequence 26, Application US/08946026
; Patent No. 6034218
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Mitcham, Jennifer L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington

; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/946.026
; FILING DATE: 07-OCT-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.424C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-946-026-26

Query Match 26.2%; Score 43; DB 3; Length 78;
Best Local Similarity 22.9%; Pred. No. 28;
Matches 8; Conservative 6; Mismatches 9; Indels 12; Gaps 1;

Qy 2 RCINQLCKLPTDSELAP-----RSXCC 24
Db 28 QCLEILCEFFQKDELKPAVTXLLWERATEKVACC 62

RESULT 18
US-08-815-718-2
; Sequence 2, Application US/08815718
; Patent No. 5981220
; GENERAL INFORMATION:
; APPLICANT: NI, JIAN
; APPLICANT: FENG, PING
; APPLICANT: DILLON, PATRICK
; APPLICANT: GENTZ, REINER
; TITLE OF INVENTION: EPIDERMAL DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
; ADDRESSEE: STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NJ
; COUNTRY: US
; ZIP: 07069
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/815.718
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 994-1700
; TELEFAX: (201) 994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-815-718-2

Query Match 26.2%; Score 43; DB 2; Length 415;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 SXCCHCR 27
| | | | |
DB 218 SRCCHCR 224

RESULT 19

US-08-884-072-3
; Sequence 3, Application US/08884072
; Patent No. 5872234
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/884,072
FILING DATE: Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 540 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAITUT13
CLONE: 1621777

US-08-884-072-3

Query Match 26.2%; Score 43; DB 2; Length 540;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 SXCCHCR 27
| | | | |
DB 218 SRCCHCR 224

RESULT 20

US-09-212-168-3
; Sequence 3, Application US/09212168
; Patent No. 6303765
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/212,168
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/884,072
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 540 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAITUT13
CLONE: 1621777
US-09-212-168-3

Query Match 26.2%; Score 43; DB 4; Length 540;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 SXCCHCR 27
| | | | |
DB 218 SRCCHCR 224

RESULT 21

US-08-946-026-3
; Sequence 3, Application US/08946026
; Patent No. 6034218
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Mitcham, Jennifer L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; AND IMMUNODIAGNOSIS OF PROSTATE CANCER
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue

Query Match 26.2%; Score 43; DB 2; Length 540;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 SXCCHCR 27
| | | | |
DB 218 SRCCHCR 224

Db 4 APRPGCHCR 13

RESULT 24

US-08-871-355A-485
; Sequence 485, Application US/08871355A
; Patent No. 6015669
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,355A
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101 CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 485:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-871-355A-485

Query Match 25.6%; Score 42; DB 3; Length 21;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 18 APRXCHCR 27

Db 4 APRPGCHCR 13

RESULT 25

US-09-201-945-485
; Sequence 485, Application US/09201945
; Patent No. 6342215
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia

COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/201,945
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/637,759
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 485:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-09-201-945-485

Query Match 25.6%; Score 42; DB 4; Length 21;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 18 APRXCHCR 27

Db 4 APRPGCHCR 13

RESULT 26

US-07-965-674-9
; Sequence 9, Application US/07965674
; Patent No. 5380646
; GENERAL INFORMATION:
; APPLICANT: Knight, Linda C.
; TITLE OF INVENTION: Thrombus Detection Using
; TITLE OF INVENTION: Radiolabelled Disintegrins
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Temple University - Of The Common-
; ADDRESSEE: wealth System of Higher Education
; STREET: 406 University Services Building
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/965,674
; FILING DATE: 19921019
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.

; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-173
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5380646e
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 62 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; US-07-965-674-9

Query Match 25.6%; Score 42; DB 1; Length 62;
Best Local Similarity 29.6%; Pred. No. 30;
Matches 8; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

QY 3 CINOLCKLPTDSELAPRSXCCHRL 29
| : ||| | : | : | :
Db 7 CCDAATCKLTPGSGCAEGLCCDNCKFI 33

RESULT 27
PCT-US93-09523-9
; Sequence 9, Application PC/TUS9309523
; GENERAL INFORMATION:
; APPLICANT: Temple University - Of The Commonwealth
; APPLICANT: System of Higher Education
; APPLICANT: Knight, Linda C. and
; APPLICANT: Maurer, Alan H.
; TITLE OF INVENTION: Thrombus Detection Using
; TITLE OF INVENTION: Radiolabelled Disintegrins
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Temple University - Of The
; ADDRESSEE: Commonwealth System of Higher Education
; STREET: 406 University Services Building
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/09523
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 965,674
; FILING DATE: 19 October 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-173 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: None
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 62 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; PCT-US93-09523-9

Query Match 25.6%; Score 42; DB 5; Length 62;
Best Local Similarity 29.6%; Pred. No. 30;
Matches 8; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

QY 3 CINOLCKLPTDSELAPRSXCCHRL 29
| : ||| | : | : | :
Db 7 CCDAATCKLTPGSGCAEGLCCDNCKFI 33

RESULT 28
5318899-16
; Patent No. 5318899
; APPLICANT: SCARBOROUGH, ROBERT M.; WOLF, DAVID L.; CHARO,
; ISRAEL F.
; TITLE OF INVENTION: PLATELET AGGREGATION INHIBITORS
; NUMBER OF SEQUENCES: 83
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/483,229
; FILING DATE: 20-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 418,028
; FILING DATE: 06-OCT-1989
; APPLICATION NUMBER: 367,509
; FILING DATE: 16-JUN-1989
; SEQ ID NO: 16:
; LENGTH: 70
; 5318899-16

Query Match 25.6%; Score 42; DB 6; Length 70;
Best Local Similarity 31.8%; Pred. No. 34;
Matches 7; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 3 CINOLCKLPTDSELAPRSXCC 24
| : ||| | : | : | :
Db 12 CCDAATCKLPPGAQCDGGGLCC 33

RESULT 29
US-09-524-162-2
; Sequence 2, Application US/09524162
; Patent No. 6355452
; GENERAL INFORMATION:
; APPLICANT: Ping Tsui
; TITLE OF INVENTION: HUMAN HISTAMINE H3 GENE VARIANT-2
; FILE REFERENCE: Gp-70681
; CURRENT APPLICATION NUMBER: US/09/524,162
; CURRENT FILING DATE: 2000-03-13
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2:
; LENGTH: 351
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
; US-09-524-162-2

Query Match 25.6%; Score 42; DB 4; Length 351;
Best Local Similarity 34.6%; Pred. No. 1.7e+02;
Matches 9; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 1 KRCINOLCKLPTDSELAPRSXCCHC 26
| : ||| | : | : | :
Db 326 RRAFTKLLC--POKLKIQPHSSLEHC 349

RESULT 30
PCT-US91-06418-1
; Sequence 1, Application PC/TUS9106418
; GENERAL INFORMATION:
; APPLICANT: Oklahoma Medical Research, Foundation, et al
; TITLE OF INVENTION: Antigens Associated with Polymyositis
; TITLE OF INVENTION: and with Dermatomyositis
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody

STREET: 100 Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/06418
FILING DATE: 19910905
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/579023
FILING DATE: 09-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRF120
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-572-6508
TELEFAX: 404-572-6555
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 351 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien
TISSUE TYPE: Sera
IMMEDIATE SOURCE:
CLONE: L1
PCT-US91-06418-1

Query Match 25.6%; Score 42; DB 5; Length 351;
Best Local Similarity 38.1%; Pred. No. 1.7e+02;
Matches 8; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 3 C1NOLLKLPDSELA PRSXC 23
I:| | | | | | | | | |
Db 186 CLNPPLPDIPNGEWICPRCTC 206

RESULT 31
US-08-985-090-2
Sequence 2, Application US/08985090
Patent No. 5885893
GENERAL INFORMATION:
APPLICANT: Andrew D. J. Goodearl
TITLE OF INVENTION: MUSCARINIC RECEPTORS AND USES THEREFOR
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/985,090

FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Jean M. Silveri
REGISTRATION NUMBER: 39,030
REFERENCE/DOCKET NUMBER: MNI-032
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 445 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-985-090-2

Query Match 25.6%; Score 42; DB 2; Length 445;
Best Local Similarity 34.6%; Pred. No. 2.1e+02;
Matches 9; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 1 KRCINOLLKLPDSELA PRSXCCHC 26
:| :||| | :| | |
Db 420 RRAFTKLLC--PQKLTQPHSSLEHC 443

RESULT 32
US-09-165-543-2
Sequence 2, Application US/09165543
Patent No. 6093545
GENERAL INFORMATION:
APPLICANT: Andrew D. J. Goodearl and Sandra Glucksmann
TITLE OF INVENTION: Muscarinic Receptors and Uses Therefor
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/165,543
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/042,780
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth A. Hanley
REGISTRATION NUMBER: 33,505
REFERENCE/DOCKET NUMBER: MNI-032CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 445 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-165-543-2

Query Match 25.6%; Score 42; DB 3; Length 445;

```
Best Local Similarity 34.6%; Pred. No. 2.1e+02;
Matches 9; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 1 KRCINQLCKLPTDSELAPRSXCCHC 26
DB 420 RRAFTKLLC--PQKLIKQPHSSLEHC 443

RESULT 33
US-09-167-354-7
; Sequence 7, Application US/09167354A
; Patent No. 6136559
; GENERAL INFORMATION:
; APPLICANT: Lovenberg, Timothy
; APPLICANT: Erlander, Mark
; APPLICANT: Pyati, Jayashree
; APPLICANT: Huvar, Arne
; TITLE OF INVENTION: DNA ENCODING A HUMAN HISTAMINE RECEPTOR OF THE H3
; TITLE OF INVENTION: SUBTYPE
; FILE REFERENCE: JWV
; CURRENT APPLICATION NUMBER: US/09/167,354A
; CURRENT FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:PEPTIDE
US-09-167-354-7

Query Match 25.6%; Score 42; DB 4; Length 445;
Best Local Similarity 34.6%; Pred. No. 2.1e+02;
Matches 9; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 1 KRCINQLCKLPTDSELAPRSXCCHC 26
DB 420 RRAFTKLLC--PQKLIKQPHSSLEHC 443

RESULT 34
US-08-759-581B-15
; Sequence 15, Application US/08759581B
; Patent No. 5876945
; GENERAL INFORMATION:
; APPLICANT: CHISHOLM, DEXTER A.
; APPLICANT: DINER, BRUCE A.
; APPLICANT: DONALDSON, GAIL K.
; APPLICANT: HERSHEY, HOWARD P.
; APPLICANT: JORDAN, DOUGLAS B.
; APPLICANT: TANG, XIAO-SONG
; APPLICANT: TROST, JEFFREY T.
; APPLICANT: WANG, SHAOJIE
; APPLICANT: WARREN, PATRICK V.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING HERBICIDAL AGENTS THAT INHIBIT D1 PROTEASE
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E.I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: U.S.A.
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 3.1
; SOFTWARE: MICROSOFT WORD 2.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/759,581B
; FILING DATE:

Best Local Similarity 34.6%; Pred. No. 2.1e+02;
Matches 9; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 1 KRCINQLCKLPTDSELAPRSXCCHC 26
DB 420 RRAFTKLLC--PQKLIKQPHSSLEHC 443

RESULT 35
US-09-304-711-15
; Sequence 15, Application US/09304711
; Patent No. 6350573
; GENERAL INFORMATION:
; APPLICANT: DINER, BRUCE A.
; APPLICANT: CHISHOLM, DEXTER A.
; APPLICANT: DONALDSON, GAIL K.
; APPLICANT: HERSHEY, HOWARD P.
; APPLICANT: JORDAN, DOUG B.
; APPLICANT: TANG, XIAO SONG
; APPLICANT: WANG, SHAOJIE
; APPLICANT: TROST, JEFFREY T.
; APPLICANT: WARREN, PATRICK V.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING HERBICIDAL AGENTS THAT INHIBIT D1 PROTEASE
; FILE REFERENCE: CR-9964-C
; CURRENT APPLICATION NUMBER: US/09/304,711
; CURRENT FILING DATE: 1999-05-04
; EARLIER APPLICATION NUMBER: 08/759,581
; EARLIER FILING DATE: DECEMBER 5, 1996
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: MICROSOFT OFFICE 97
; SEQ ID NO 15
; LENGTH: 473
; TYPE: PRT
; ORGANISM: TOBACCO
US-09-304-711-15

Query Match 25.6%; Score 42; DB 4; Length 473;
Best Local Similarity 41.2%; Pred. No. 2.2e+02;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 4 INQLCKLPTDSELAPR 20
DB 262 VKSRICKLPTGGDDAPQ 278

RESULT 36
US-09-049-672A-1
; Sequence 1, Application US/09049672A
; Patent No. 6135941
; GENERAL INFORMATION:
```

```

; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836.443
; FILING DATE: 01-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9709420
; FILING DATE: 05-SEP-1997
; APPLICATION NUMBER: 9612145.4
; FILING DATE: 11-JUN-1996
; APPLICATION NUMBER: 9526230.9
; FILING DATE: 21-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 47424
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 751 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5883241e
; US-08-836-443-3

Query Match 25.6%; Score 42; DB 2; Length 751;
Best Local Similarity 39.3%; Pred. No. 3.5e+02;
Matches 11; Conservative 2; Mismatches 13; Indels 2; Gaps 2;

QY 3 CINOLLCKLPYDSELAPRSXCCH-CRHL 29
| | | | | | | | | | | | | | | |
Db 267 CCNASNCTLRPGARCA-HGSCCHQCKLL 293

RESULT 38
US-09-027-337-2
; Sequence 2, Application US/09027337B
; Patent No. 5972616
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotoshi
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease Overexpressed in
; FILE OF INVENTION: Breast and Ovarian Carcinomas
; FILE REFERENCE: D6064
; CURRENT APPLICATION NUMBER: US/09/027.337B
; CURRENT FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 13
; SEQ ID NO 2
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of TADG-15 encoded by nucleotides
; OTHER INFORMATION: 23 to 2589 of Sequence 1
; Patent No. 5972616
; US-09-027-337-2

Query Match 25.6%; Score 42; DB 2; Length 855;
Best Local Similarity 36.4%; Pred. No. 4e+02;
Matches 8; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

```

QY 2 RCINQLCKLPTDSELAPRSXC 23
| | | | | : | | |
Db 769 RVINQTTCCNLLPQOITPRMC 790

RESULT 39

US-09-113-750A-3
; Sequence 3, Application US/09113750A
; Patent No. 6294176

; GENERAL INFORMATION:

; APPLICANT: David E. Junker and Mark D. Cochran
; TITLE OF INVENTION: Recombinant Raccoonpox virus

; NUMBER OF SEQUENCES: 58

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: John P. White

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/113,750A

; FILING DATE:

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 55744

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212)262-0400

; TELEFAX: (212)664-0525

; TELEX: 422523

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 688 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-09-113-750A-3

Query Match 25.3%; Score 41.5; DB 4; Length 688;

Best Local Similarity 38.5%; Pred. No. 3.8e+02;

Matches 10; Conservative 2; Mismatches 7; Indels 7; Gaps 1;

QY 2 RCIN-----QLLCKLPTDSELAPR 20
| | | | | : | | | | |

Db 483 RCINFLYRELKSVNCLDGDGLDR 508

RESULT 40

US-09-813-819-4

; Sequence 4, Application US/09813819

; Patent No. 6294368

; GENERAL INFORMATION:

; APPLICANT: MERKULOV, Gennady et al

; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND

; FILE OF INVENTION: USES THEREOF

; FILE REFERENCE: CL001177

; CURRENT APPLICATION NUMBER: US/09/813,819

; CURRENT FILING DATE: 2001-03-22

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 814
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(814)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-813-819-4

Query Match 25.3%; Score 41.5; DB 4; Length 814;

Best Local Similarity 30.8%; Pred. No. 4.5e+02;

Matches 8; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

QY 3 CINQLCKLPTDSELAPRSKCC-HCR 27
| : | | | : | | | |

Db 447 CCDSLTCQLRPGAQACASDGGPCQNCQ 472

Search completed: August 22, 2002, 07:50:49

Job time: 187 sec